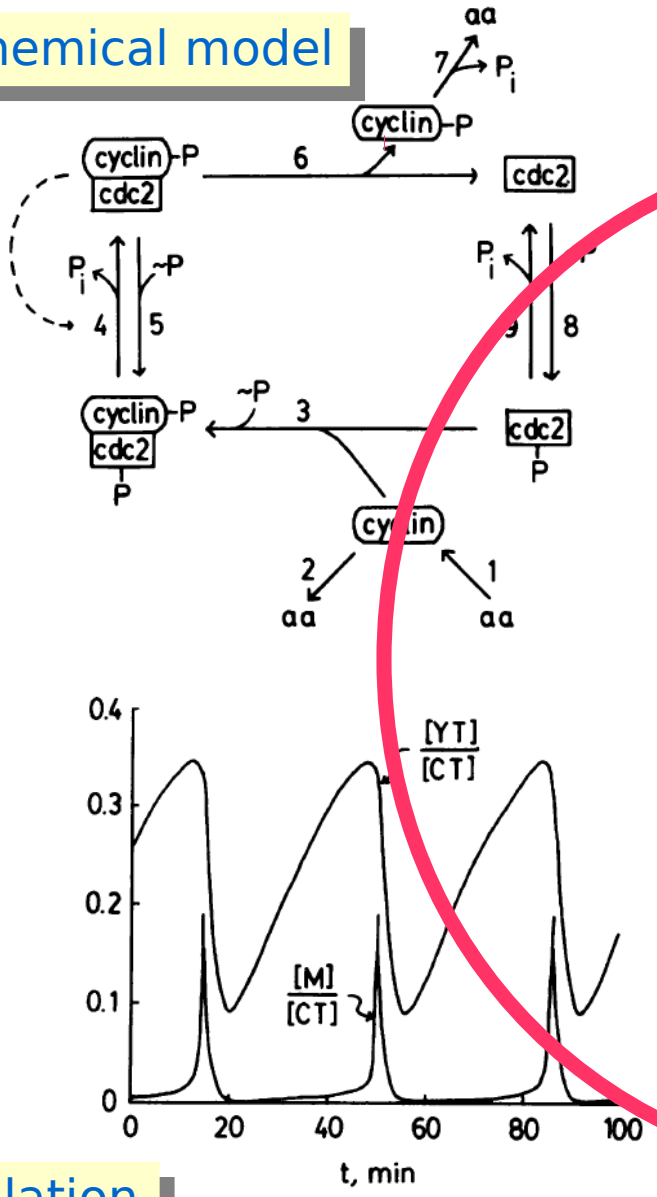
Knowledge representation and ontologies in Systems Biology

Nicolas Le Novère, EMBL-EBI

The models I am going to talk about

biochemical model



simulation

mathematical model

$$d[C2]/dt = k_6[M] - k_8[\sim P][C2] + k_9[CP]$$
$$d[CP]/dt = -k_3[CP][Y] + k_8[\sim P][C2] - k_9[CP]$$
$$d[pM]/dt = k_3[CP][Y] - [pM]F([M]) + k_5[\sim P][M]$$
$$d[M]/dt = [pM]F([M]) - k_5[\sim P][M] - k_6[M]$$
$$d[Y]/dt = k_1[aa] - k_2[Y] - k_3[CP][Y]$$
$$d[YP]/dt = k_6[M] - k_7[YP]$$

Parameter	Value	Notes
$k_1[aa]/[CT]$	0.015 min^{-1}	*
k_2	0	†
$k_3[CT]$	200 min^{-1}	*
k_4	10–1000 min^{-1} (adjustable)	
k_4'	0.018 min^{-1}	
$k_5[\sim P]$	0	‡
k_6	0.1–10 min^{-1} (adjustable)	
k_7	0.6 min^{-1}	†
$k_8[\sim P]$	$>> k_9$	§
k_9	$>> k_6$	§

computational model

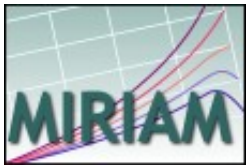
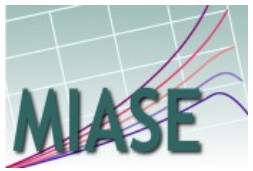
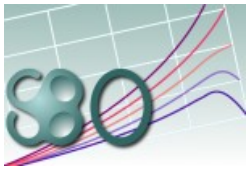
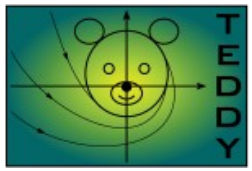
Tyson et al (1991) PNAS 88(1):7328-32

Models at the core of integrative systems biology

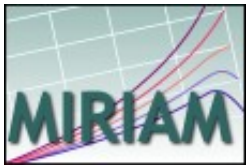
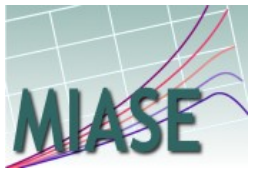
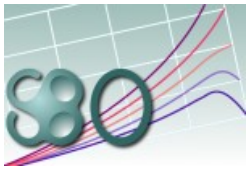
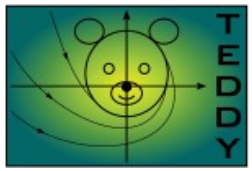
- Computational models encode the relationships between the building blocks of a system
- Computational models can be linked to experimental datasets: controlled links
- Computational models can be exchanged: computer storable and readable
- Computational models can be analysed in many different ways by different tools: standard API
- Computational models can be modified, split, merged: Encoded in a suitable design

➔ Standardisation is unavoidable!

A “complete” (?) mosaic of standards

	Models	Simulation	Results
Minimal requirements			
Data-models		SED-ML	SBRML
Ontologies			

Model description

	Models	Simulation	Results
Minimal requirements			
Data-model		ED-ML	SBRML
Ontologies			

The Systems Biology Markup Language

News Documents Downloads Forums Facilities Community Events About

Google Site Search..

The **Systems Biology Markup Language** (**SBML**) is a computer-readable format for representing models of biological processes. It's applicable to simulations of metabolism, cell-signaling, and many other topics. SBML has been evolving since mid-2000 thanks to an international community of software developers and users. This website is the global portal for the SBML effort; here you can find information about all aspects of SBML.

For the curious

What *is* SBML? Read our **basic introduction** and then perhaps browse the **mailing lists** to get a sense for what's currently going on in the world of SBML.

For modelers

Are you looking for ready-to-run software that supports SBML? The **SBML Software Guide** lists over **180** systems today. Are you instead looking for ready-to-use models? Visit the **BioModels Database**, where you can find hundreds!

For software developers

Are you interested in developing SBML support for your software? Read our **basic introduction** and then the **SBML specifications** to understand how to use SBML. After that, you may want to look at **libSBML**, an API library supporting many programming languages.

No matter how you use SBML, we invite you to sign up for news updates either through our **RSS feed**, our **Twitter feed**, or one of the **mailing lists**, and get involved with community efforts to help keep improving SBML. You can also call

SBML News

SBML-BioModels.net Hackathon

(25 Jan.'10) The next Hackathon will be held May 1–4 in Seattle. Please **let us know** if you're coming.

LibSBML 4.0.1 released!

(21 Jan.'10) This release of the libSBML version 4.x series fixes many reported small bugs in 4.0.0.

LibSBML 5 preview!

(18 Jan.'10) A preview version of libSBML version 5, with a plug-in architecture for Level 3 packages, is **now available**.

Older news ...

Community News

BioModels release 16

(26 Jan.'09) The latest release of the BioModels Database features over 450 total models.

Cain 1.2 released!

(1 Jan.'09) **Cain** is a stochastic simulator with highly efficient implementations of many methods.

BioUML Web Edition

(30 Nov.'09) BioUML, the Java-based modeling platform, now has a **web edition**.

Hucka M., H. Bolouri, A. Finney, H. M. Sauro, J. C. Doyle, H. Kitano et al (2003). The Systems Biology Markup Language (SBML): A Medium for Representation and Exchange of Biochemical Network Models. *Bioinformatics*, 19: 524-531.

What is SBML?

The Systems Biology Markup Language is a way to **exchange and reuse** (and hopefully **interface**) descriptions of quantitative models in “Systems Biology”, in fact mostly well-stirred chemical kinetics so far.

It is not a procedural language.

It is not a programming language.

It is not a format for specific software configuration files.

Tools are supposed to understand the whole of SBML but not obligatory make use of everything (will not be true for Level 3).

Development philosophy: Start small, get good support, extend.

SBML itself re-uses other standards: MathML, XHTML, RDF, existing ontologies.

It is supported by a community large, diverse, active and **evolving.**

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="4" xmlns="http://www.sbml.org/sbml/level2/version4">
  <model>
    <listOfCompartments>
      <compartment id="cell" />
    </listOfCompartments>
    <listOfSpecies>
      <species id="A" compartment="cell" initialConcentration="1"/>
      <species id="B" compartment="cell" initialConcentration="0"/>
    </listOfSpecies>
    <listOfParameters>
      <parameter id="kon" value="1"/>
    </listOfParameters>
    <listOfReactions>
      <reaction>
        <listOfReactants>
          <speciesReference species="A" />
        </listOfReactants>
        <listOfProducts>
          <speciesReference species="B" />
        </listOfProducts>
        <kineticLaw>
          <math xmlns="http://www.w3.org/1998/Math/MathML">
            <apply>
              <times />
              <ci>kon</ci>
              <ci>A</ci>
              <ci>cell</ci>
            </apply>
          </math>
        </kineticLaw>
      </reaction>
    </listOfReactions>
  </model>
</sbml>
```

No biological semantics in the language itself

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="4" xmlns="http://www.sbml.org/sbml/level2/version4">
  <model>
    <listOfCompartments>
      <compartment id="cell" />
    </listOfCompartments>
    <listOfSpecies>
      <species id="A" compartment="cell" initialConcentration="1"/>
      <species id="B" compartment="cell" initialConcentration="0"/>
    </listOfSpecies>
    <listOfParameters>
      <parameter id="kon" value="1"/>
    </listOfParameters>
    <listOfReactions>
      <reaction>
        <listOfReactants>
          <speciesReference species="A" />
        </listOfReactants>
        <listOfProducts>
          <speciesReference species="B" />
        </listOfProducts>
        <kineticLaw>
          <math xmlns="http://www.w3.org/1998/Math/MathML">
            <apply>
              <times />
              <ci>kon</ci>
              <ci>A</ci>
              <ci>cell</ci>
            </apply>
          </math>
        </kineticLaw>
      </reaction>
    </listOfReactions>
  </model>
</sbml>
```

Elements free of biological semantics: we do not know which type of species this is

Reusing existing standards

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="1" xmlns="http://www.sbml.org/sbml/level2">
  <model>
    <listOfCompartments>
      <compartment id="cell" />
    </listOfCompartments>
    <listOfSpecies>
      <species id="A" compartment="cell" initialConcentration="1"/>
      <species id="B" compartment="cell" initialConcentration="0"/>
    </listOfSpecies>
    <listOfParameters>
      <parameter id="kon" value="1"/>
    </listOfParameters>
    <listOfReactions>
      <reaction>
        <listOfReactants>
          <speciesReference species="A" />
        </listOfReactants>
        <listOfProducts>
          <speciesReference species="B" />
        </listOfProducts>
        <kineticLaw>
          <math xmlns="http://www.w3.org/1998/Math/MathML">
            <apply>
              <times />
              <ci>kon</ci>
              <ci>A</ci>
              <ci>cell</ci>
            </apply>
          </math>
        </kineticLaw>
      </reaction>
    </listOfReactions>
  </model>
</sbml>
```

MathML

A more realistic example ...

```
<species
    id="A"
    name="α-tubulin"
    compartment="cell"
    initialAmount="1000"
    substanceUnits="item"
    hasOnlySubstanceUnits="true"
    boundaryCondition="true"
    constant="false"
    charge="0"
    metaid="PX"
    sboTerm="SBO:0000245" >
  <notes>
    <body xmlns="http://www.w3.org/1999/xhtml">
      <p>One of the components of a microtubule</p>
    </body>
  </notes>
  <annotation>
    <rdf:RDF
        xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
        xmlns:bqmodel="http://biomodels.net/model-qualifiers/"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
      <rdf:Description rdf:about="#PX">
        <bqbiol:is>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:uniprot:P68370"/>
            <rdf:li rdf:resource="urn:miriam:obo.go:GO%3A0045298"/>
          </rdf:Bag>
        </bqbiol:is>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

Externalisation of biological semantics

```
<species
    id="A"
    name="α-tubulin"
    compartment="cell"
    initialAmount="1000"
    substanceUnits="item"
    hasOnlySubstanceUnits="true"
    boundaryCondition="true"
    constant="false"
    charge="0"
    metaid="PX"
    sboTerm="SBO:0000245" >
  <notes>
    <body xmlns="http://www.w3.org/1999/xhtml">
      <p>One of the components of a microtubule</p>
    </body>
  </notes>
  <annotation>
    <rdf:RDF
        xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
        xmlns:bqmodel="http://biomodels.net/model-qualifiers/"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
      <rdf:Description rdf:about="#PX">
        <bqbiol:is>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:uniprot:P68370"/>
            <rdf:li rdf:resource="urn:miriam:obo.go:GO%3A0045298"/>
          </rdf:Bag>
        </bqbiol:is>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

macromolecule

Reusing existing standards again

```
<species
    id="A"
    name="α-tubulin"
    compartment="cell"
    initialAmount="1000"
    substanceUnits="item"
    hasOnlySubstanceUnits="true"
    boundaryCondition="true"
    constant="false"
    charge="0"
    metaid="PX"
    sboTerm="SBO:0000245" >
  <notes>
    <body xmlns="http://www.w3.org/1999/xhtml">
      <p>One of the components of a microtubule</p>
    </body>
  </notes>
  <annotation>
    <rdf:RDF
        xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
        xmlns:bqmodel="http://biomodels.net/model-qualifiers/"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
      <rdf:Description rdf:about="#PX">
        <bqbiol:is>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:uniprot:P68370"/>
            <rdf:li rdf:resource="urn:miriam:obo.go:GO%3A0045298"/>
          </rdf:Bag>
        </bqbiol:is>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

XHTML

RDF

SBML is not limited to biochemistry!

Rate Rules can describe the temporal evolution of any quantitative parameter, e.g. transmembrane voltage;

Events can describe any discontinuous change, e.g. neurotransmitter release or repolarisation;

A species is an entity participating to a reaction, **not always** a **chemical** entity:

- It can be a molecule
- It can be a cell
- It can be an organ
- It can be an organism

➔ Systems Biology is scale-free!

Hodgkin-Huxley

```
<rateRule metaid="metaid_0000048" variable="V">
  <notes><p xmlns="http://www.w3.org/1999/xhtml">dV/dt = (I - (i_Na + i_K + i_L))/Cm</p></notes>
  <math xmlns="http://www.w3.org/1998/Math/MathML">
    <apply>
      <divide/>
      <apply>
        <minus/>
        <ci> I </ci>
        <apply>
          <plus/><ci> i_Na </ci><ci> i_K </ci><ci> i_L </ci>
        </apply>
      </apply>
      <ci> Cm </ci>
    </apply>
  </math>
</rateRule>
<assignmentRule metaid="metaid_0000042" variable="i_Na">
  <notes><p xmlns="http://www.w3.org/1999/xhtml">i_Na = g_Na * m^3.0 * h * (V - E_Na)</p></notes>
  <math xmlns="http://www.w3.org/1998/Math/MathML">
    <apply>
      <times/>
      <ci> g_Na </ci>
      <apply>
        <power/><ci> m </ci><cn> 3.0 </cn>
      </apply>
      <ci> h </ci>
      <apply>
        <minus/><ci> V </ci><ci> E_Na </ci>
      </apply>
    </apply>
  </math>
</assignmentRule>
```

rate rule:

dx/dt = f (x,y,z)

assignment rule:

x = f (y,z)

BioModels Home | Browse models | Submit | Search

http://www.ebi.ac.uk/biomodels/

BIOMD0000000020 - Hodgkin-Huxley1952 squid-axon

SBML L2 V1 | Other formats | Actions | Submit Model Comment/Bug

Model | Overview | Math | Physical entities | Parameters

Reference Publication

Publication ID: 12991237

J Physiol 1952 Aug;117(4):500-44.
A quantitative description of membrane current and its application to conduction and excitation in nerve.
HODGKIN AL, HUXLEY AF.
[more]

Model

Original Model: *Unspecified*

Submitter: Nicolas Le Novère

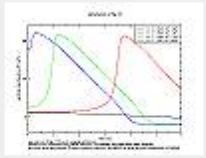
Submission Date: 2005-09-13T13:22:40+00:00

Last Modification Date: 2009-03-05T23:37:11+00:00

Creation Date: 2005-03-31T13:09:21+00:00

Creators: Maria Schilstra
Catherine Lloyd

Curation Result:

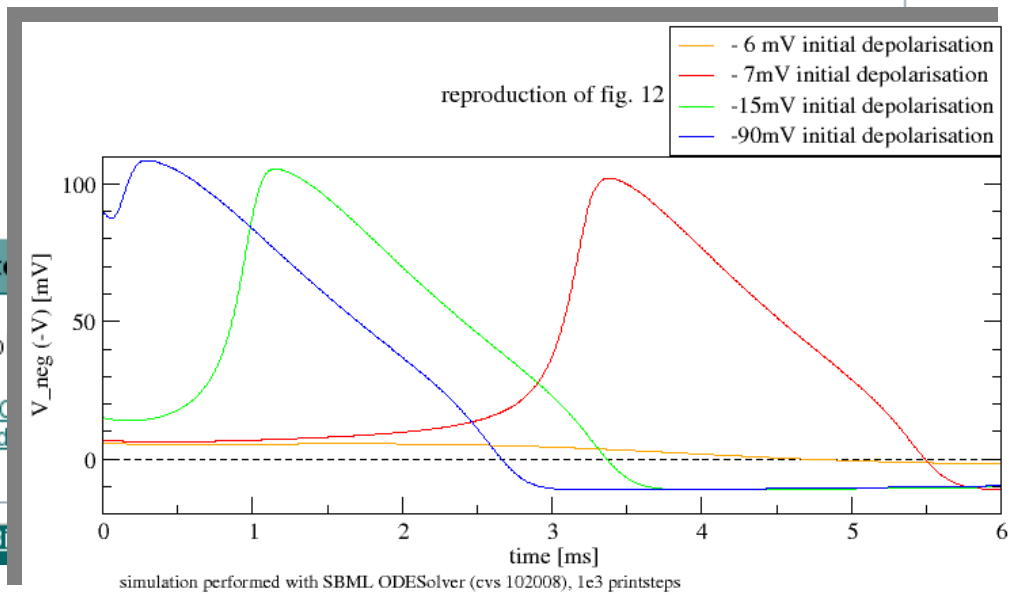


set #1	bqbiol:is	Taxonomy Loligo forbesi
set #2	bqbiol:isVersionOf	Gene Ontology voltage-gated potassium channel activity Gene Ontology voltage-gated sodium channel activity Gene Ontology action potential propagation

Not

This model originates from BioModels Database: A Database of Annotated Pub
For more information see the terms of use.
To cite BioModels Database, please use Le Novère N., Bornstein B., Broicher A., C
J.L., Hucka M. (2006) BioModels Database: A Free, Centralized Database of Curated
Nucleic Acids Res., 34: D689-D691.

Developed by BioModels Team of Computational Neurobiology Group in European Bi



An example of pharmacodynamic model

```
<rateRule metaid="metaid_0000031" variable="Size">
  <math xmlns="http://www.w3.org/1998/Math/MathML">
    <apply>
      <times/>
      <apply>
        <minus/>
        <apply>
          <times/><ci> RateIn </ci><ci> Effect </ci>
        </apply>
        <apply>
          <times/><ci> Kover </ci><ci> Size </ci>
        </apply>
      </apply>
      <ci> Size </ci>
    </apply>
  </math>
</rateRule>
<assignmentRule metaid="metaid_0000027" variable="Effect">
  <math xmlns="http://www.w3.org/1998/Math/MathML">
    <apply>
      <minus/>
      <cn> 1 </cn>
      <apply>
        <divide/>
        <ci> Ce </ci>
        <apply>
          <plus/><ci> AE50 </ci><ci> Ce </ci>
        </apply>
      </apply>
    </apply>
  </math>
</assignmentRule>
```

Tham et al (2008) A pharmacodynamic model for the time course of tumor shrinkage by gemcitabine + carboplatin in non-small cell lung cancer patients. Clin Cancer Res. 2008 14(13): 4213-8.

$$\frac{Size}{dt} = (Rate_{in} \times Effect - K_{over} \times Size) \times Size$$

$$Effect = 1 - \frac{E_{max} - Ce}{Amt_{50} + Ce}$$

BIOMD0000000234 - Tham2008_PDmode_TumourShrinkage

SBML formats | Other formats | Actions | Curation Comment

Model | Overview | Math | Physical entities | Parameters | Curation

Reference Publication

Publication ID: 18594002

Clin Cancer Res 2008 Jul;14(13):4213-8.
A pharmacodynamic model for the time course of tumor shrinkage by gemcitabine + carboplatin in non-small cell lung cancer patients.
Tham LS, Wang L, Soo RA, Lee SC, Lee HS, Yong WP, Goh BC, Holford NH.
Department of Hematology-Oncology, National University Hospital, Singapore. Tham_lai_san@lilly.com [more]

Model

Original Model: *BIOMD0000000234.xml.origin*

Submitter: Nick Holford

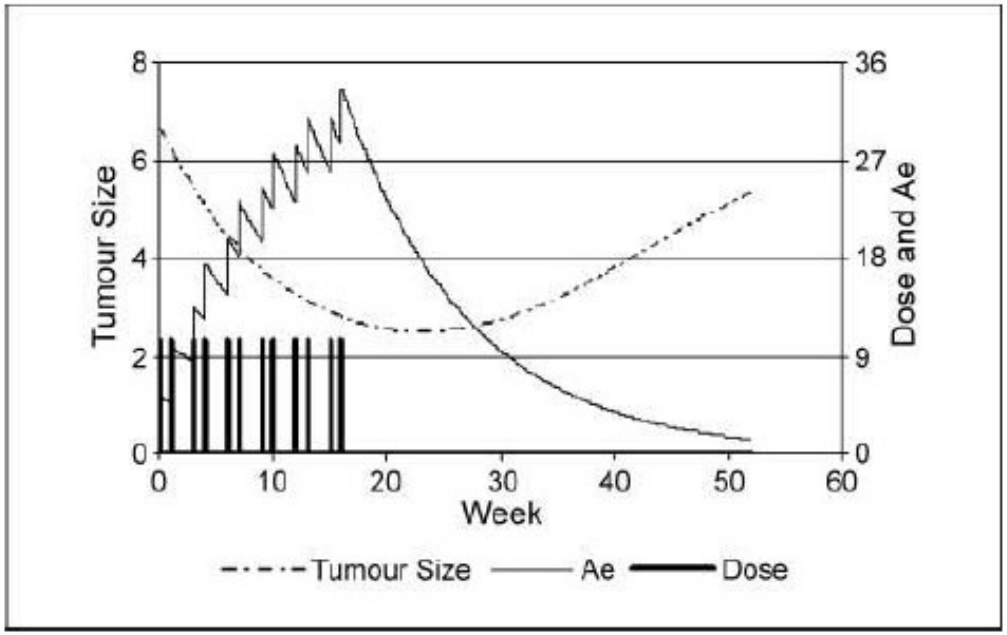
Submission ID: MODEL0911120001

Submission Date: 2009-11-16T19:41:59+00:00

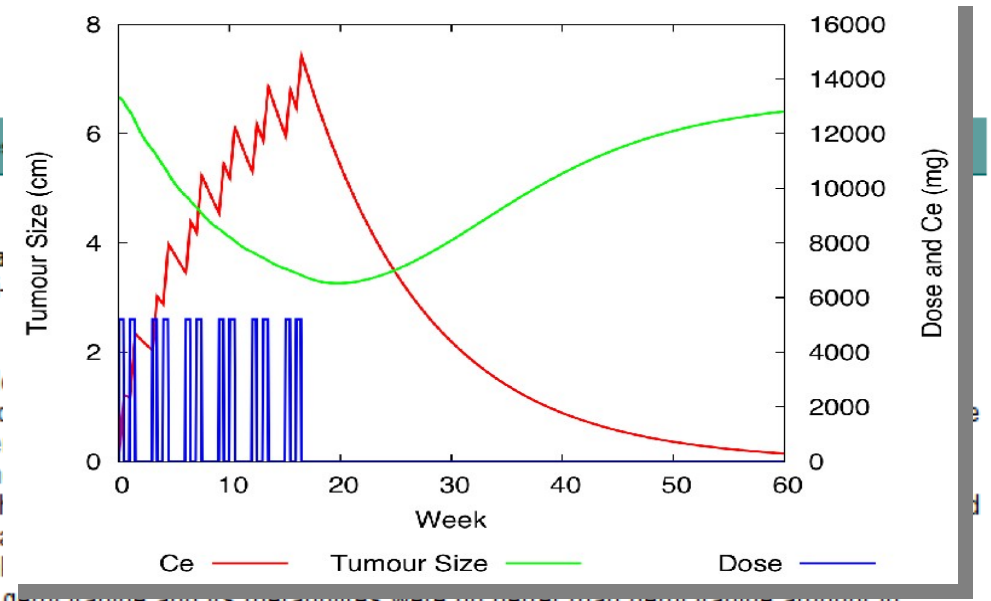
Last Modification Date: 2009-11-23T13:24:05+00:00

set #1	bqbiol:occursIn	Taxonomy Homo sapiens
set #2	bqbiol:isVersionOf	ICD C34 OMIM 211980

original article



reproduction with SBMLodeSolver



predicting tumor shrinkage in primary lung cancer lesions. Gemcitabine dose-based models did marginally better than treatment-based models that ignored doses of drug administered to patients. Modeling tumor shrinkage in primary lesions can be used to quantify individual sensitivity and response to antitumor effects of anticancer

Difference between SBML L1, L2 and L3

L1

- predefined functions
- proprietary infix math notation
- reserved namespaces for annotation
- no controlled annotation
- no discrete events
- monolithic
- default values

L2

- function definitions
- all math in MathML
- no reserved namespaces for annotations
- controlled RDF annotation
- discrete events
- monolithic
- default values

L3

- function definitions
- all math in MathML
- no reserved namespaces for annotations
- controlled RDF annotation
- discrete events
- modular
- no default values

Progressive simplification, generalisation and externalisation

A glimpse of SBML Level 3

Modular SBML, with core + optional packages

- Core package – Release candidate
- Graph Layout – specification finalised
- Complex species – specification finalised
- Groups - specification under discussion
- Model composition – specification under discussion
- Qualitative models – specification under discussion
- Graph rendering – specification proposed
- Distributions and ranges - specification proposed
- Arrays and sets – specifications proposed
- Geometry - specification proposed
- Spatial diffusion – needed
- Dynamic structures - needed

???

SBML Software Guide

SBML software and projects come in many varieties. Here we summarize all SBML-compatible systems known to us. The *matrix* provides an at-a-glance summary, whereas the *summary* provides longer descriptions of each software or project grouped by themes. Please **use the survey form** to notify us about additions and suggestions.

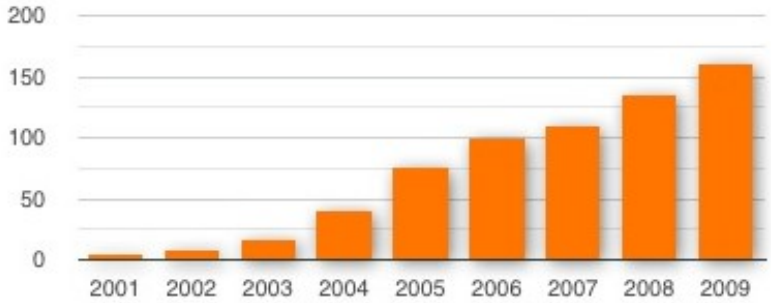
Number of software packages listed in the matrix today: **181**

Go to the SBML Software Matrix

Go to the SBML Software Summary

Historical trend

The following graph shows the total number of known SBML-compatible software packages each year, as counted by the SBML Team. The counts shown are for approximately the middle of each year.

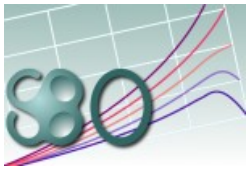
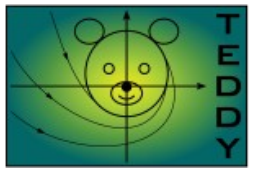


(Note: the flat period in 2007 is an artifact of inadequate record keeping rather than a lull in SBML software development.)

Disclaimers

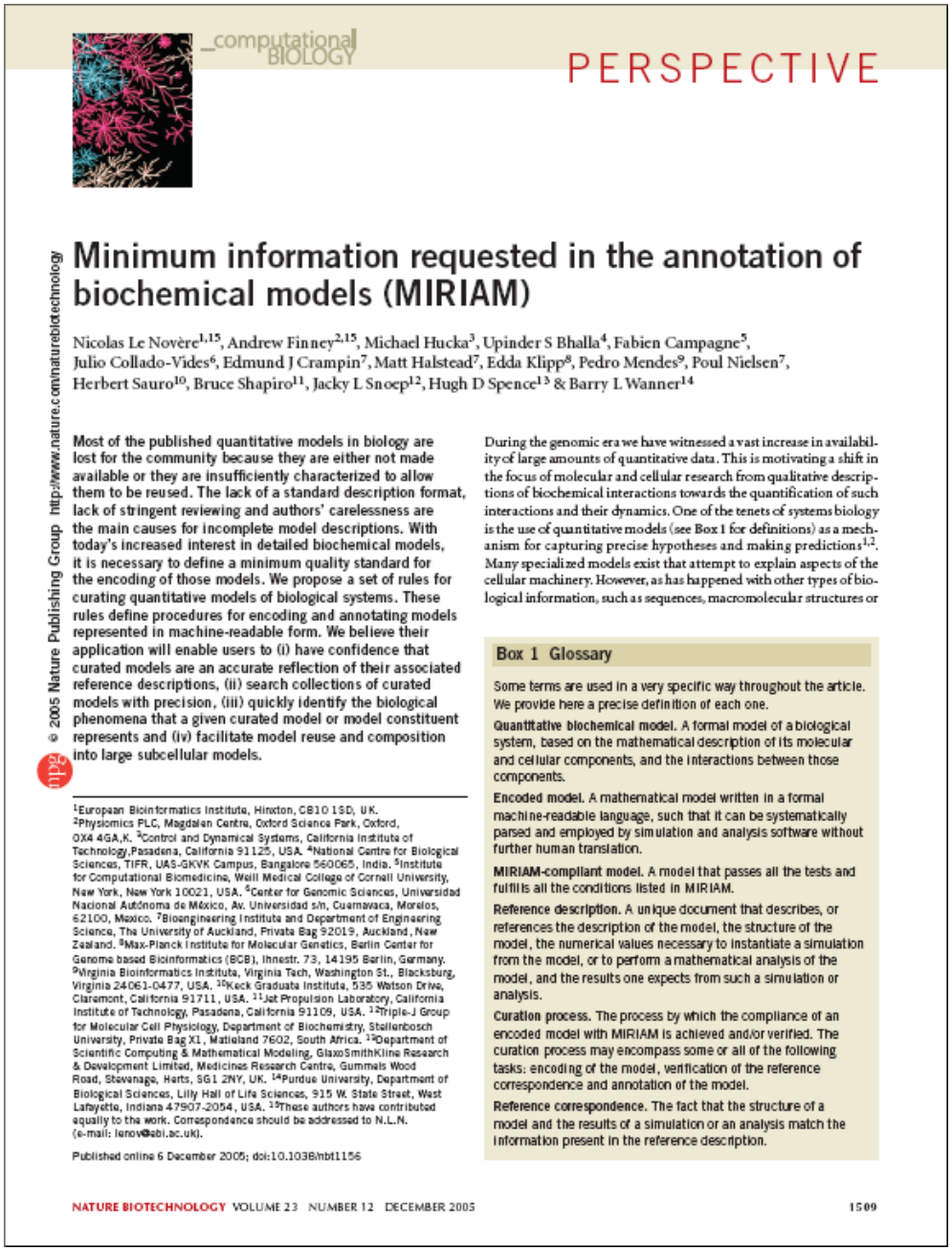
The listing of a software project or other work on this page does not constitute an endorsement from anyone associated with the SBML Project or the institutions supporting the SBML Project. The SBML Team makes a best-faith effort to provide accurate

Model semantics

	Models	Simulation	Results
Minimal requirement			
Data-models	cellML biology. math. data. knowledge. SBML SBGN	SED-ML	SBRML
Ontologies			

MIRIAM guidelines

- Reporting guidelines for curation of quantitative models
  - Specifically about encoding & annotation
  - Limited for the moment to models that can be numerically evaluated
- Not specific to SBML; applicable to any structured model format



computational BIOLOGY

PERSPECTIVE

Minimum information requested in the annotation of biochemical models (MIRIAM)

Nicolas Le Novère[1,15], Andrew Finney[2,15], Michael Hucka[3], Upinder S Bhalla[4], Fabien Campagne[5], Julio Collado-Vides[6], Edmund J Crampin[7], Matt Halstead[7], Edda Klipp[8], Pedro Mendes[9], Poul Nielsen[7], Herbert Sauro[10], Bruce Shapiro[11], Jacky L Snoep[12], Hugh D Spence[13] & Barry L Wanner[14]

Most of the published quantitative models in biology are lost for the community because they are either not made available or they are insufficiently characterized to allow them to be reused. The lack of a standard description format, lack of stringent reviewing and authors' carelessness are the main causes for incomplete model descriptions. With today's increased interest in detailed biochemical models, it is necessary to define a minimum quality standard for the encoding of those models. We propose a set of rules for curating quantitative models of biological systems. These rules define procedures for encoding and annotating models represented in machine-readable form. We believe their application will enable users to (i) have confidence that curated models are an accurate reflection of their associated reference descriptions, (ii) search collections of curated models with precision, (iii) quickly identify the biological phenomena that a given curated model or model constituent represents and (iv) facilitate model reuse and composition into large subcellular models.

[1]European Bioinformatics Institute, Hinxton, CB10 1SD, UK. [2]Physiomics PLC, Magdalen Centre, Oxford Science Park, Oxford, OX4 4GA,K. [3]Control and Dynamical Systems, California Institute of Technology,Pasadena, California 91125, USA. [4]National Centre for Biological Sciences, TIFR, UAS-GKVK Campus, Bangalore 560065, India. [5]Institute for Computational Biomedicine, Weill Medical College of Cornell University, New York, New York 10021, USA. [6]Center for Genomic Sciences, Universidad Nacional Autónoma de México, Av. Universidad s/n, Cuernavaca, Morelos, 62100, Mexico. [7]Bioengineering Institute and Department of Engineering Science, The University of Auckland, Private Bag 92019, Auckland, New Zealand. [8]Max-Planck Institute for Molecular Genetics, Berlin Center for Genome based Bioinformatics (BCB), Ihnestr. 73, 14195 Berlin, Germany. [9]Virginia Bioinformatics Institute, Virginia Tech, Washington St., Blacksburg, Virginia 24061-0477, USA. [10]Keck Graduate Institute, 535 Watson Drive, Claremont, California 91711, USA. [11]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA. [12]Triple-J Group for Molecular Cell Physiology, Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa. [13]Department of Scientific Computing & Mathematical Modeling, GlaxoSmithKline Research & Development Limited, Medicines Research Centre, Gunnels Wood Road, Stevenage, Herts, SG1 2NY, UK. [14]Purdue University, Department of Biological Sciences, Lilly Hall of Life Sciences, 915 W. State Street, West Lafayette, Indiana 47907-2054, USA. [15]These authors have contributed equally to the work. Correspondence should be addressed to N.L.N. (e-mail: lenov@ebi.ac.uk).

Published online 6 December 2005; doi:10.1038/nbt1156

During the genomic era we have witnessed a vast increase in availability of large amounts of quantitative data. This is motivating a shift in the focus of molecular and cellular research from qualitative descriptions of biochemical interactions towards the quantification of such interactions and their dynamics. One of the tenets of systems biology is the use of quantitative models (see Box 1 for definitions) as a mechanism for capturing precise hypotheses and making predictions[1,2]. Many specialized models exist that attempt to explain aspects of the cellular machinery. However, as has happened with other types of biological information, such as sequences, macromolecular structures or

Box 1 Glossary

Some terms are used in a very specific way throughout the article. We provide here a precise definition of each one.

Quantitative biochemical model. A formal model of a biological system, based on the mathematical description of its molecular and cellular components, and the interactions between those components.

Encoded model. A mathematical model written in a formal machine-readable language, such that it can be systematically parsed and employed by simulation and analysis software without further human translation.

MIRIAM-compliant model. A model that passes all the tests and fulfills all the conditions listed in MIRIAM.

Reference description. A unique document that describes, or references the description of the model, the structure of the model, the numerical values necessary to instantiate a simulation from the model, or to perform a mathematical analysis of the model, and the results one expects from such a simulation or analysis.

Curation process. The process by which the compliance of an encoded model with MIRIAM is achieved and/or verified. The curation process may encompass some or all of the following tasks: encoding of the model, verification of the reference correspondence and annotation of the model.

Reference correspondence. The fact that the structure of a model and the results of a simulation or an analysis match the information present in the reference description.

MIRIAM compliance

Models must :

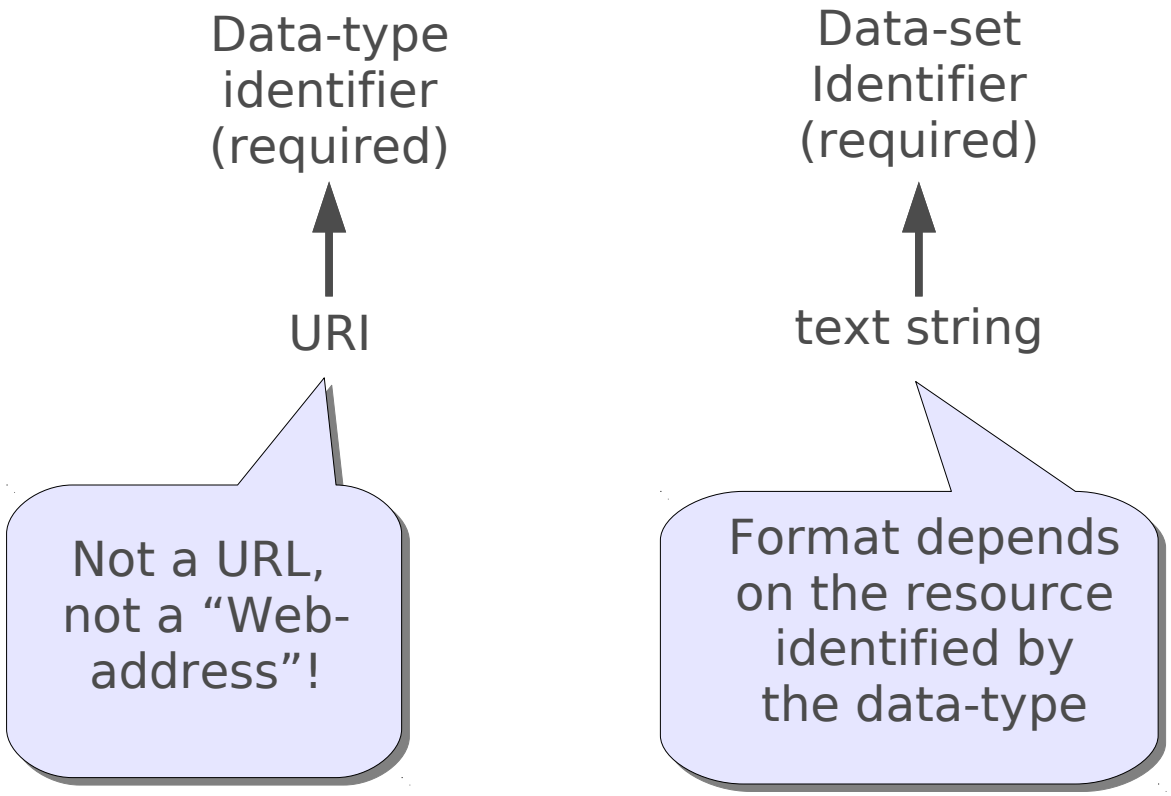
- be encoded in a public machine-readable format
- be clearly linked to a single reference description
- reflect the structure of the biological processes described in the reference paper (list of reactions etc.)
- be instantiable in a simulation (possess initial conditions etc.)
- be able to reproduce the results given in the reference paper
- contain creator's contact details
- annotation to unambiguously identify each model constituent

Why are annotations important?

Annotation of model components are essential to:

- unambiguously identify model components
  - improve understanding the structure of the model
  - allow easier comparison of different models
  - ease the integration of models
- allow efficient search strategies
- add a semantic layer to the model
  - improve understanding of the biology behind the model
  - allow conversion and reuse of the model
  - ease the integration of model and biological knowledge

MIRIAM annotation

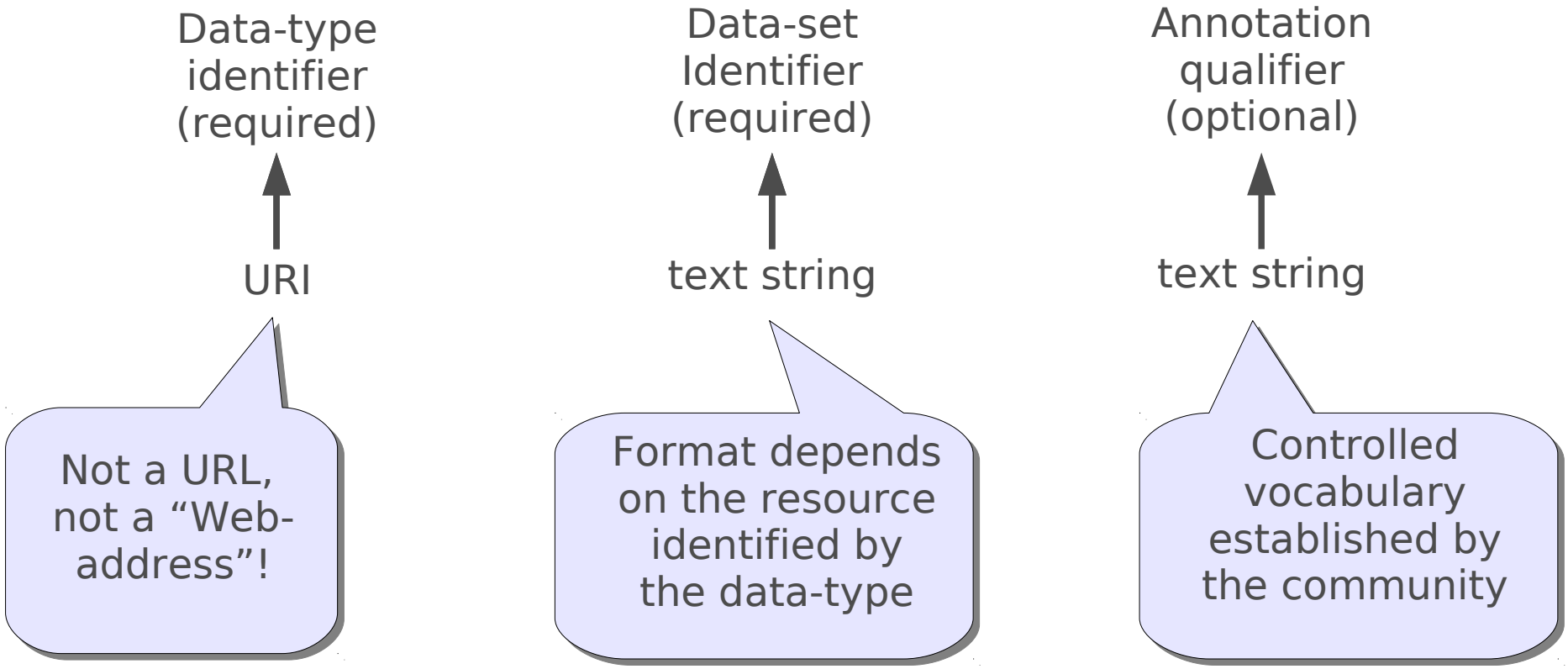


UniProt and P62158 (human calmodulin)
→ urn:miriam:uniprot:P62158

EC code and 1.1.1.1 (alcohol dehydrogenase)
→ urn:miriam:ec-code:1.1.1.1

Gene Ontology and GO:0000186 (activation of MAPKK activity)
→ urn:miriam:obo.go:GO%3A0000186

MIRIAM annotation

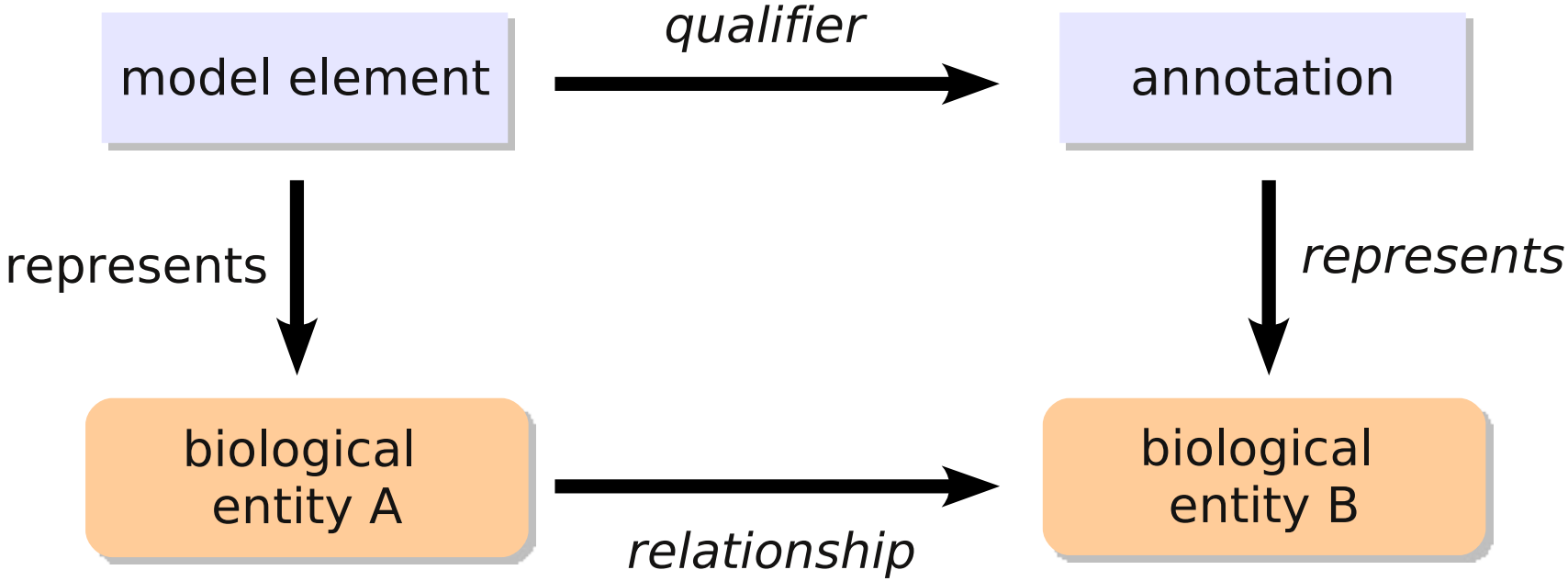


UniProt and P62158 (human calmodulin)
→ urn:miriam:uniprot:P62158

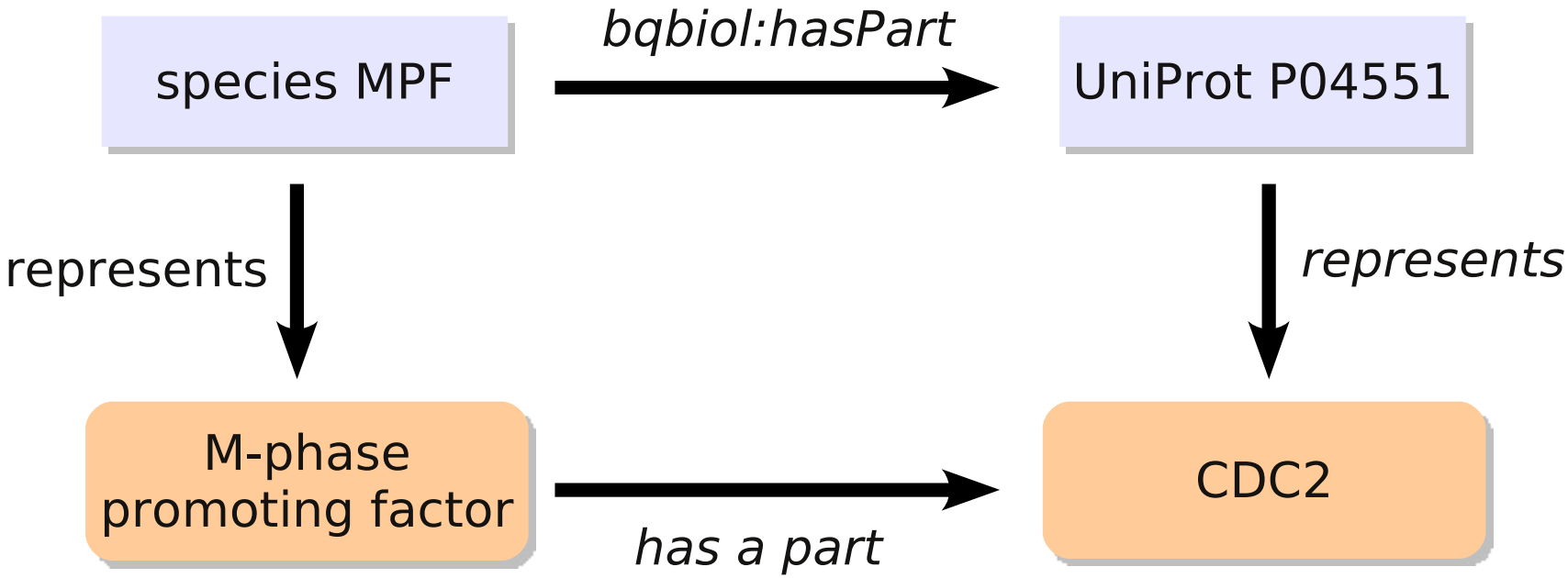
EC code and 1.1.1.1 (alcohol dehydrogenase)
→ urn:miriam:ec-code:1.1.1.1

Gene Ontology and GO:0000186 (activation of MAPKK activity)
→ urn:miriam:obo.go:GO%3A0000186

Qualification of annotation



Qualification of annotation



```
<species id="Ca_calmodulin" metaid="cacam">
  <annotation>
    <rdf:RDF
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:bqbiol="http://biomodels.net/biology-qualifiers/">
      <rdf:Description rdf:about="#cacam">
        <bqbiol:hasPart>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:uniprot:P62158"/>
            <rdf:li rdf:resource="urn:miriam:obo.chebi:CHEBI%3A29108"/>
          </rdf:Bag>
        </bqbiol:hasPart>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

Tools developing support for MIRIAM identifiers

- Data resources
  - BioModels Database (kinetic models)
  - PSI consortium (protein interactions)
  - Reactome (pathways)
  - SABIO-RK (reaction kinetics)
  - Yeast consensus model database
  - Human consensus model database
  - E-MeP (structural genomics)
- MIRIAM Resources statistics
  - ~5000 web page requests per month
  - ~550000 web service requests per month
- Application software
  - ARCADIA (graph editor)
  - BIOUML (modeling and simulation)
  - COPASI (Simulation)
  - libAnnotationSBML
  - libSBML
  - SAINT (semantic annotation)
  - SBML2BioPAX
  - SBML2LaTeX
  - SBMLeditor (model editor)
  - SemanticSBML (annotation and merging)
  - Snazer (Network analysis, Simulations)
  - Systems Biology Workbench (model design and simulation)
  - The Virtual Cell (Simulation)

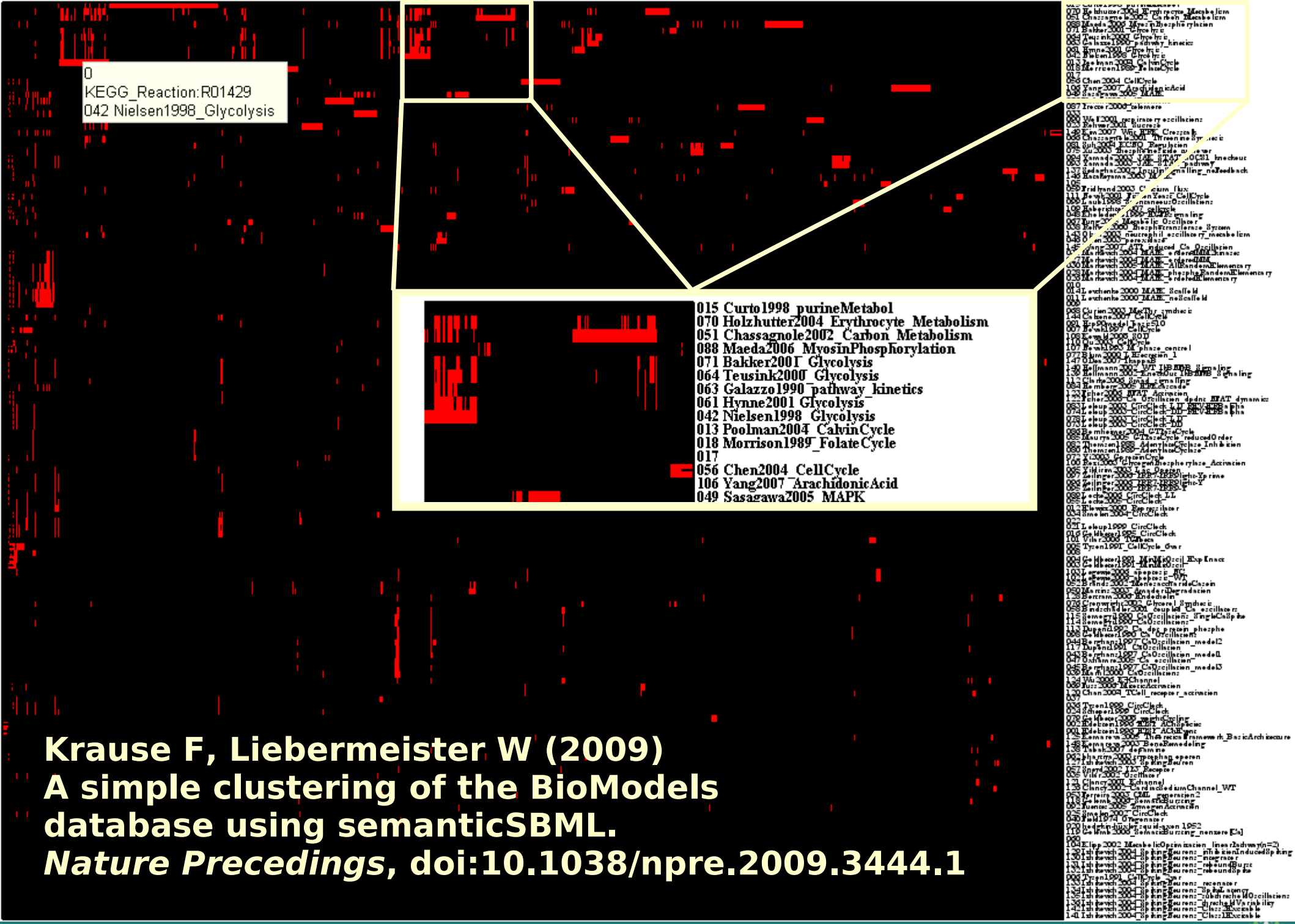
Application: whole-cell metabolic models

- Yeast Metabolic Model
  Herrgård M.J., Swainston N., Dobson P., Dunn W.B., Arga K.V., Arvas M., Blüthgen N., Borger S., Costenoble R, Heinemann M., Hucka M., Le Novère N., Li P., Liebermeister W., Mo M.L., Oliveira A.P., Petranovic D., Pettifer S., Simeonidis E., Smallbone K., Spasic I., Weichart D., Brent R., Broomhead D.S., Westerhoff H.V., Kırdar B., Penttilä M., Klipp E., Palsson B.Ø., Sauer U., Oliver S.G., Mendes P., Nielsen J., Kell D.B. (2008) A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nature Biotechnology*, 26: 1155-1160.

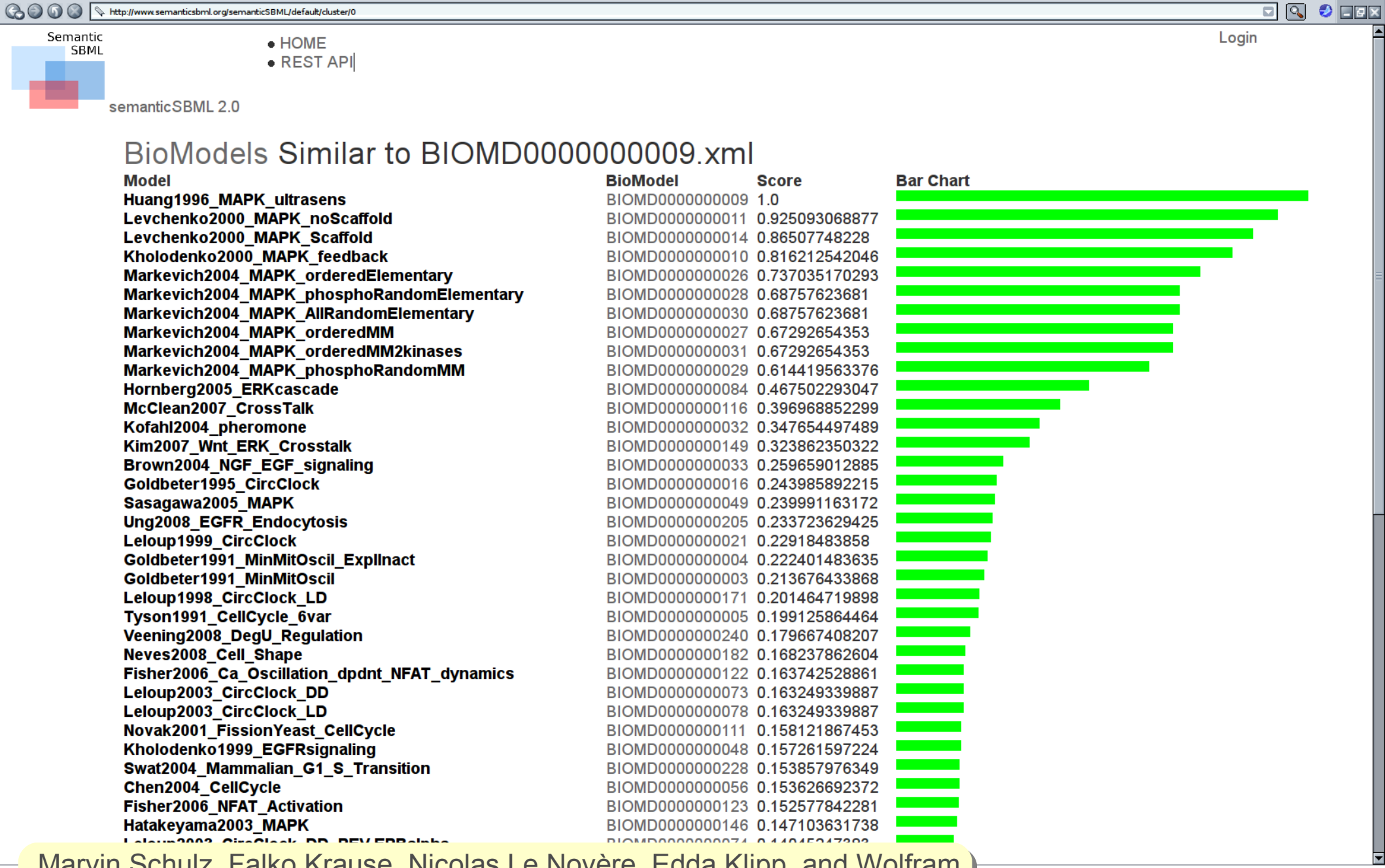
  2152 species, 1857 reactions, 4861 MIRIAM annotations

- Human Metabolic Model

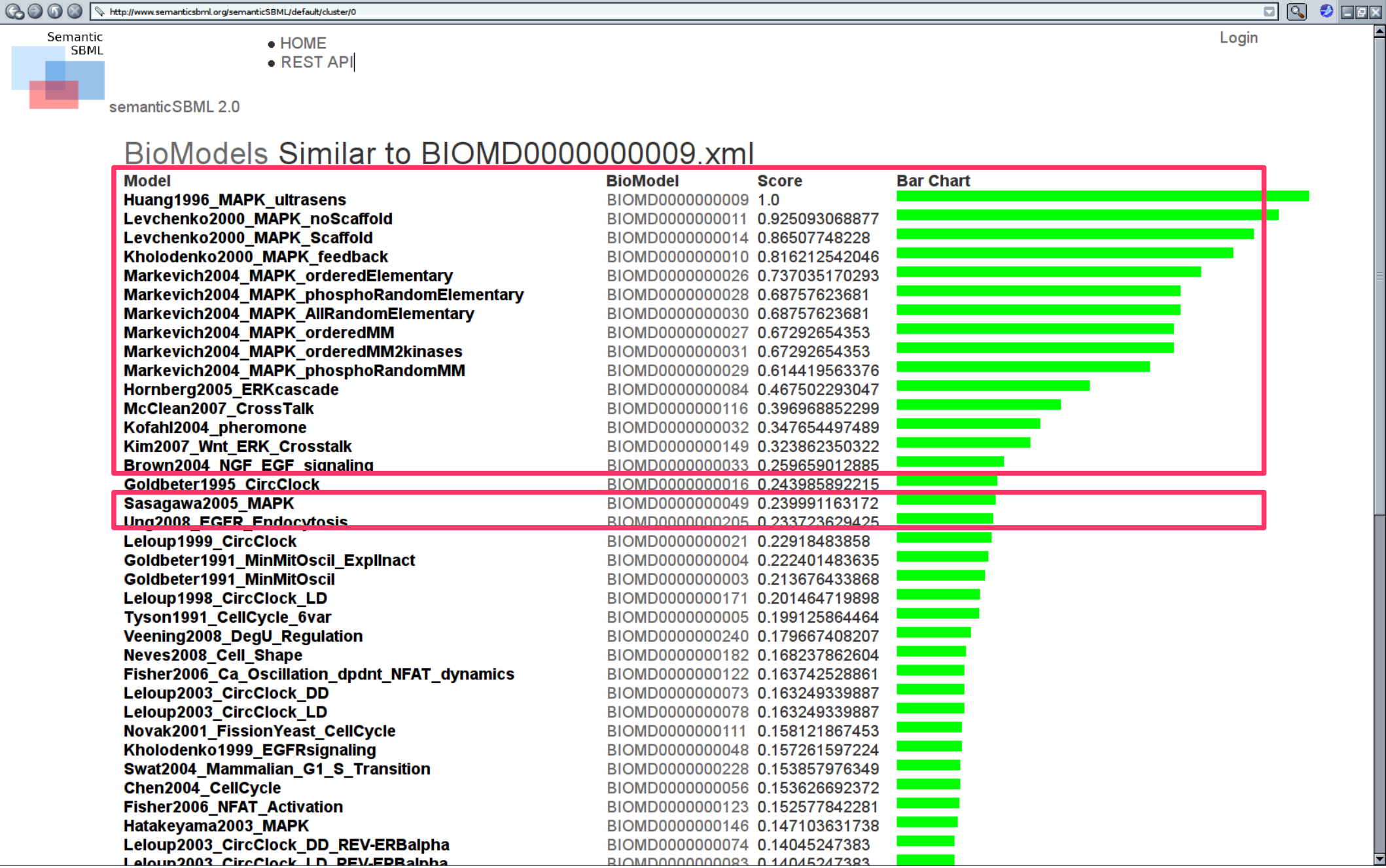
  4889 species, 8866 reactions, **66968** MIRIAM annotations



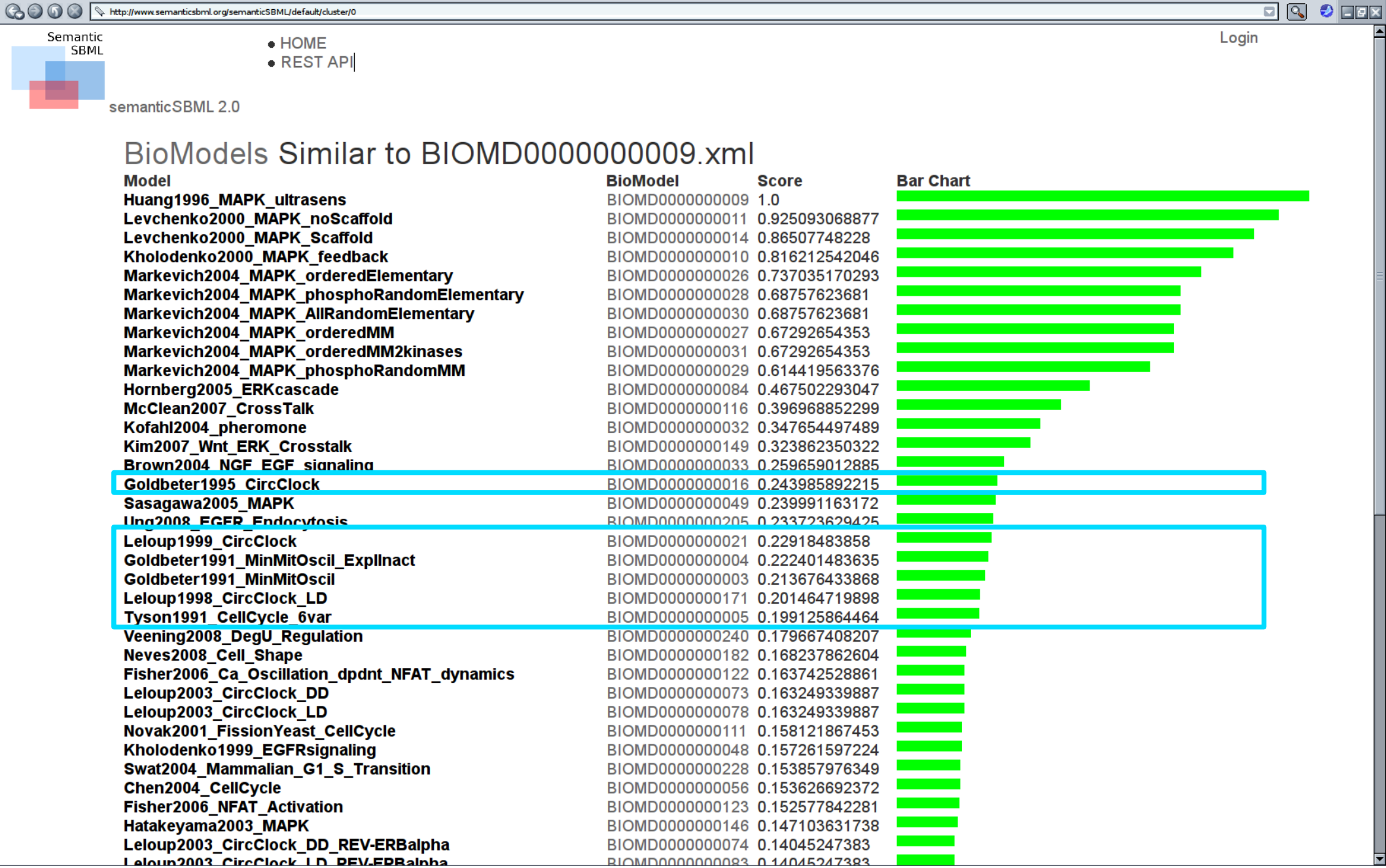
Krause F, Liebermeister W (2009)
A simple clustering of the BioModels database using semanticSBML.
***Nature Precedings*, doi:10.1038/npre.2009.3444.1**



Marvin Schulz, Falko Krause, Nicolas Le Novère, Edda Klipp, and Wolfram Liebermeister: Comparison and clustering of biochemical network models based on semantic annotations (in preparation)

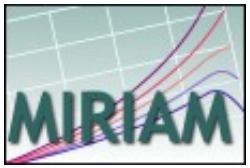
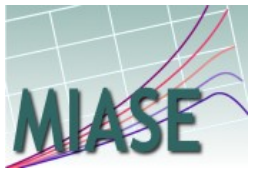
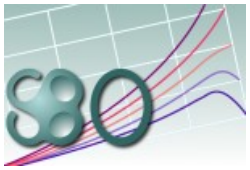
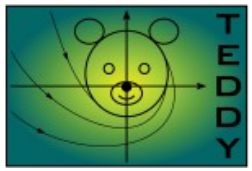


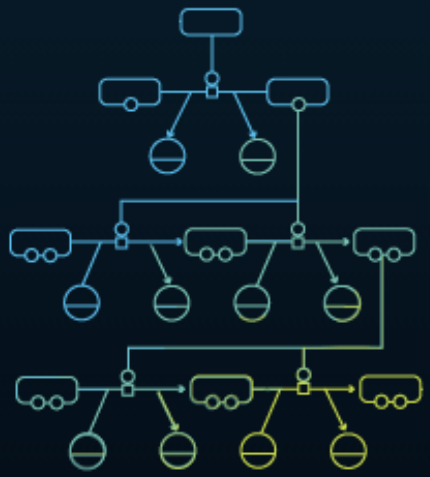
MAP Kinase cascade



Other kinase cascades

Model representation

	Models	Simulation	Results
Minimal requirements			
Data-model		ED-ML	SBRML
Ontologies			



A Visual Notation for Network Diagrams in Biology

SBGN.org is the global portal for documentation, news, and other information about the Systems Biology Graphical Notation (SBGN) project, an effort to standardize the graphical notation used in maps of biochemical and cellular processes studied in systems biology.

Standardizing the visual representation is crucial for more efficient and accurate transmission of biological knowledge between different communities in research, education, publishing, and more. When biologists are as familiar with the notation as electronics engineers are familiar with the notation of circuit schematics, they can save the time and effort required to familiarize themselves with different notations, and instead spend more time thinking about the biology being depicted.

SBGN is made up of [| three orthogonal languages], representing different visions of biological systems. Each language defines a comprehensive set of symbols with precise semantics, together with detailed syntactic rules how maps are to be interpreted.

On this site, you can browse some example maps to get a feeling for SBGN, read the SBGN specification documents , join online discussions, see current working documents and software support in the SBGN wiki, and much more.

SBGN is the work of many people. It would not have been possible without the generous support of multiple organizations over the years, for which we are very thankful.

To quote SBGN as a whole, please use:

Le Novère N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, Bergman FT, Gauges R, Ghazal P, Kawaji H, Li L, Matsuoka Y, Villéger A, Boyd SE, Calzone L, Courtot M, Dogrusoz U, Freeman TC, Funahashi A, Ghosh S, Jouraku A, Kim S, Kolpakov F, Luna A, Sahle S, Schmidt E, Watterson S, Wu G, Goryanin I, Kell DB, Sander C, Sauro H, Snoep JL, Kohn K, Kitano H. The Systems Biology Graphical Notation. Nat Biotechnol. 2009 27(8):735-41.

SBGN News

(02 Sep.'09) The first specifications for SBGN Entity Relationships and SBGN Activity Flows] are out.

What is SBGN?

- An unambiguous way of graphically describing and interpreting biochemical and cellular events
- Limited amount of symbols
  Re-use existing symbols

  ☞ Smooth learning curve
- Can represent logical or mechanistic models, biochemical pathways, at different levels of granularity
- Detailed technical specification, precise data-models and growing software support
- Initiated by Hiroaki Kitano. Developed over four years by a diverse community

The Systems Biology Graphical Notation

Nicolas Le Novère[1], Michael Hucka[2], Huaiyu Mi[3], Stuart Moodie[4], Falk Schreiber[5,6], Anatoly Sorokin[7], Emek Demir[8], Katja Wegner[9], Mirit I Aladjem[10], Sarala M Wimalaratne[11], Frank T Bergman[12], Ralph Gauges[13], Peter Ghazal[4,14], Hideya Kawaji[15], Lu Li[1], Yukiko Matsuoka[16], Alice Villéger[17,18], Sarah E Boyd[19], Laurence Calzone[20], Melanie Courtot[21], Ugur Dogrusoz[22], Tom C Freeman[14,23], Akira Funahashi[24], Samik Ghosh[16], Akiya Jouraku[24], Sohyoung Kim[10], Fedor Kolpakov[25,26], Augustin Luna[10], Sven Sahle[13], Esther Schmidt[1], Steven Watterson[4,22], Guanming Wu[27], Igor Goryanin[4], Douglas B Kell[18,28], Chris Sander[8], Herbert Sauro[12], Jacky L Snoep[29], Kurt Kohn[10] & Hiroaki Kitano[16,30,31]

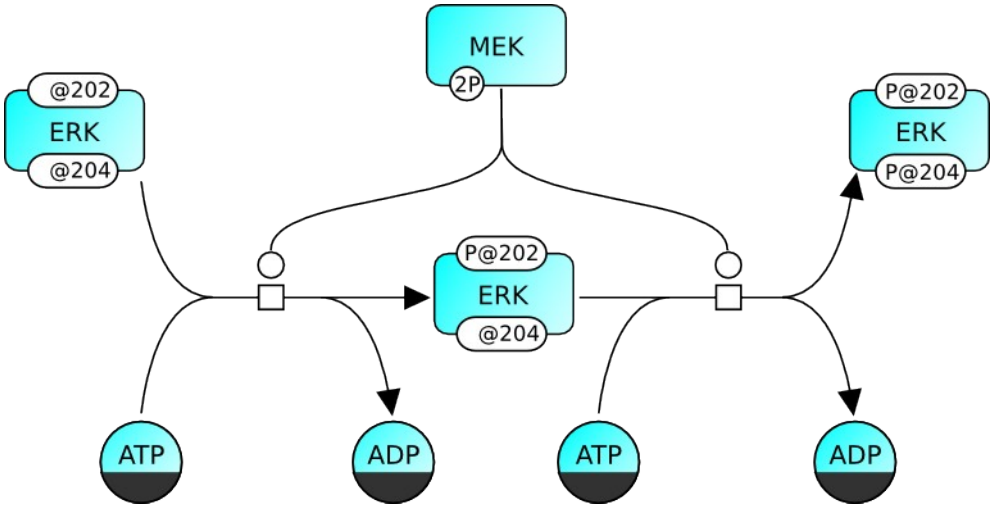
[1]EMBL European Bioinformatics Institute, Hinxton, UK. [2]Engineering and Applied Science, California Institute of Technology, Pasadena, California, USA. [3]SRI International, Menlo Park, California, USA. [4]Centre for Systems Biology at Edinburgh, University of Edinburgh, Edinburgh, UK. [5]Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany. [6]Institute of Computer Science, University of Halle, Halle, Germany. [7]School of Informatics, University of Edinburgh, Edinburgh, UK. [8]Memorial Sloan Kettering Cancer Center - Computational Biology Center, New York, NY, USA. [9]Science and Technology Research Institute, University of Hertfordshire, Hatfield, UK. [10]National Cancer Institute, Bethesda, Maryland, USA. [11]Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand. [12]Department of Bioengineering, University of Washington, Seattle, Washington, USA. [13]BIOQUANT, University of Heidelberg, Heidelberg, Germany. [14]Division of Pathway Medicine, University of Edinburgh Medical School, Edinburgh, UK. [15]Riken OMICS Science Center, Yokohama City, Kanagawa, Japan. [16]The Systems Biology Institute, Tokyo, Japan. [17]School of Computer Science, University of Manchester, Manchester, UK. [18]Manchester Interdisciplinary Biocentre, Manchester, UK. [19]Clayton School of Information Technology, Faculty of Information Technology, Monash University, Melbourne, Victoria, Australia. [20]U900 INSERM, Paris Mines Tech, Institut Curie, Paris, France. [21]Terry Fox Laboratory, British Columbia Cancer Research Center, Vancouver, British Columbia, Canada. [22]Bilkent Center for Bioinformatics, Bilkent University, Ankara, Turkey. [23]The Roslin Institute, University of Edinburgh, Midlothian, UK. [24]Department of Biosciences and Informatics, Keio University, Hiyoshi, Kouhoku-ku, Yokohama, Japan. [25]Institute of Systems Biology, Novosibirsk, Russia. [26]Design Technological Institute of Digital Techniques SB RAS, Novosibirsk, Russia. [27]Ontario Institute for Cancer Research, Toronto, Ontario, Canada. [28]School of Chemistry, University of Manchester, Manchester, UK. [29]Department of Biochemistry, Stellenbosch University, Matieland, South Africa. [30]Sony Computer Science Laboratories, Tokyo, Japan. [31]Okinawa Institute of Science and Technology, Okinawa, Japan. Correspondence should be addressed to N.L.N. (lenov@ebi.ac.uk).

NATURE BIOTECHNOLOGY VOLUME 27 NUMBER 8 AUGUST 2009

39 authors, 31 affiliations

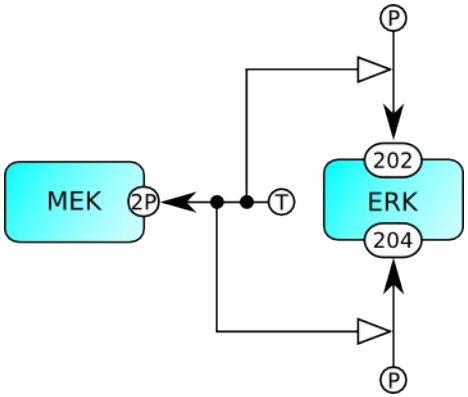
Graph trinity: three languages in one notation

Process Descriptions



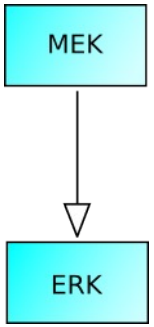
- Unambiguous
- Mechanistic
- Sequential
- Combinatorial explosion

Entity Relationships



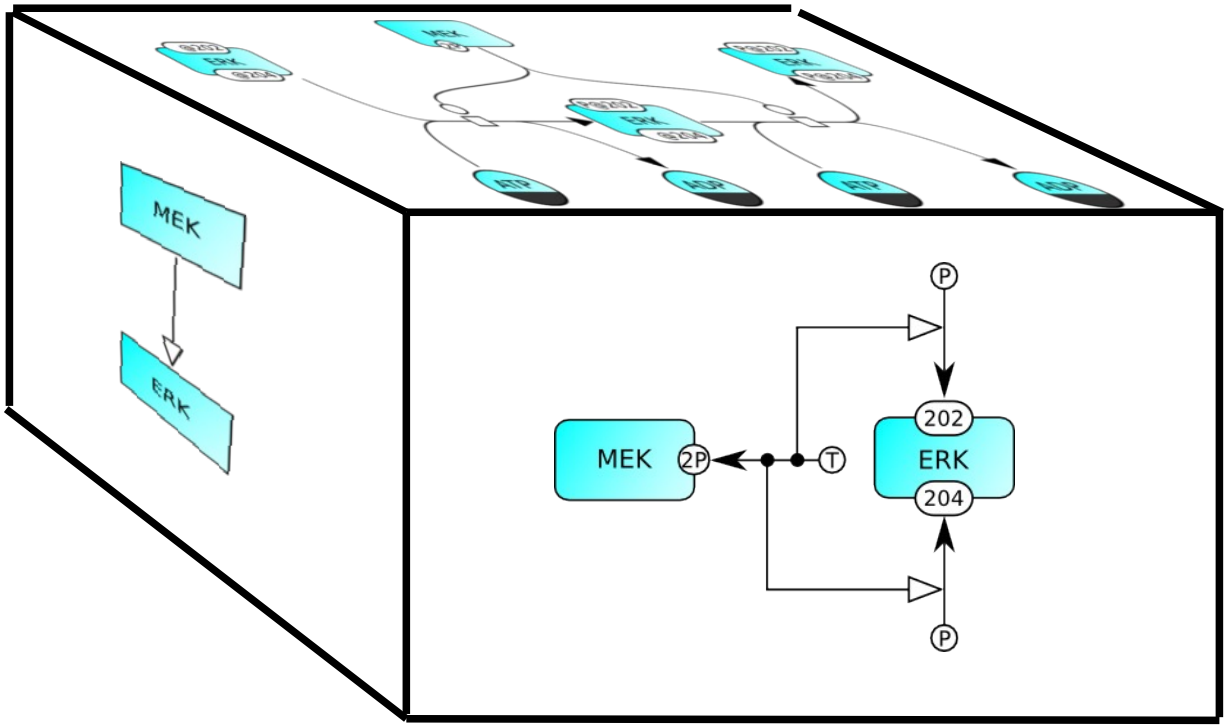
- Unambiguous
- Mechanistic
- Non-sequential
- Independence of relationships

Activity Flows

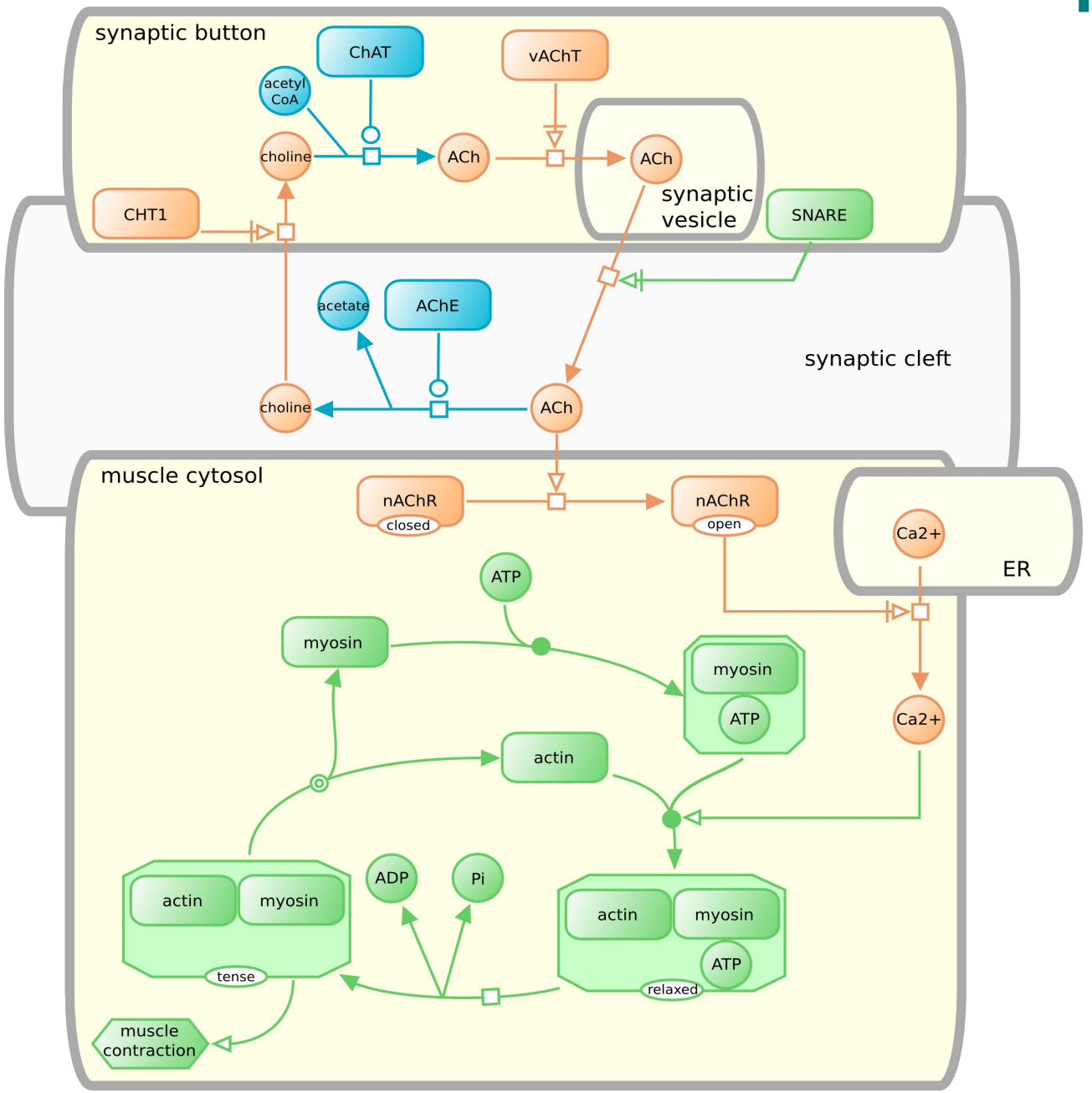


- Ambiguous
- Conceptual
- Sequential

Three orthogonal projections of biology



multi-cellular processes in Process Description language

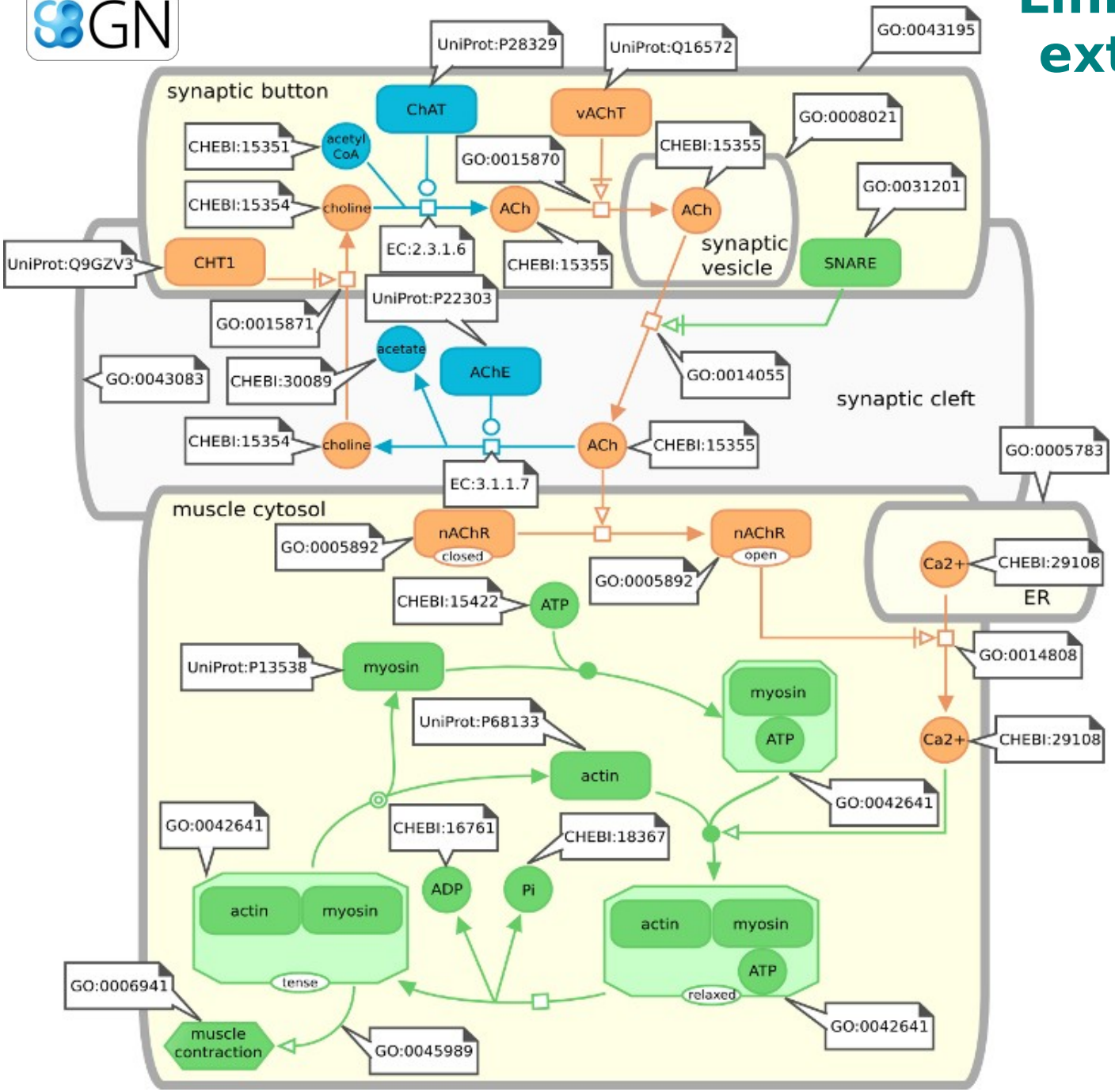


catalytic processes

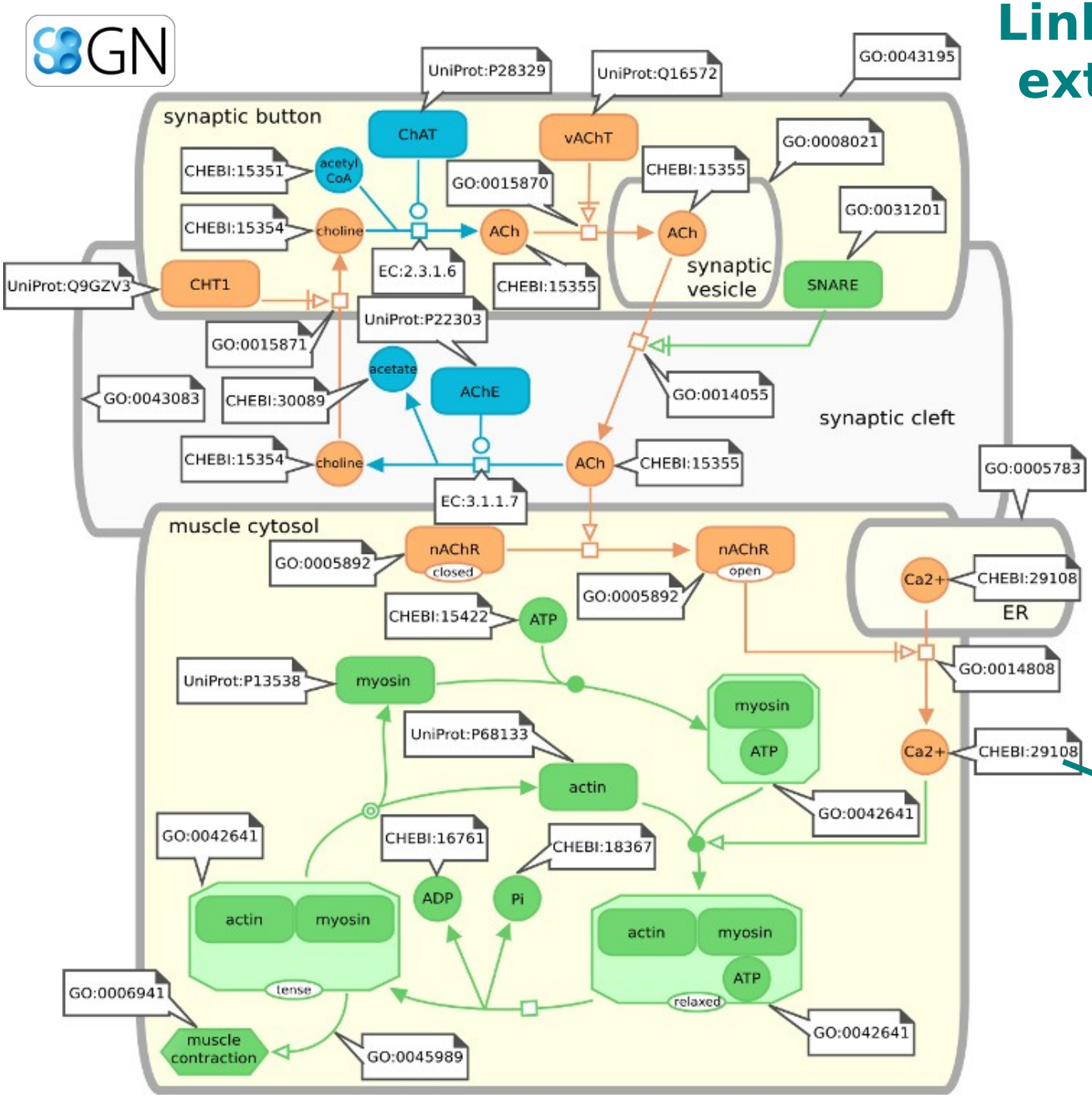
transport processes

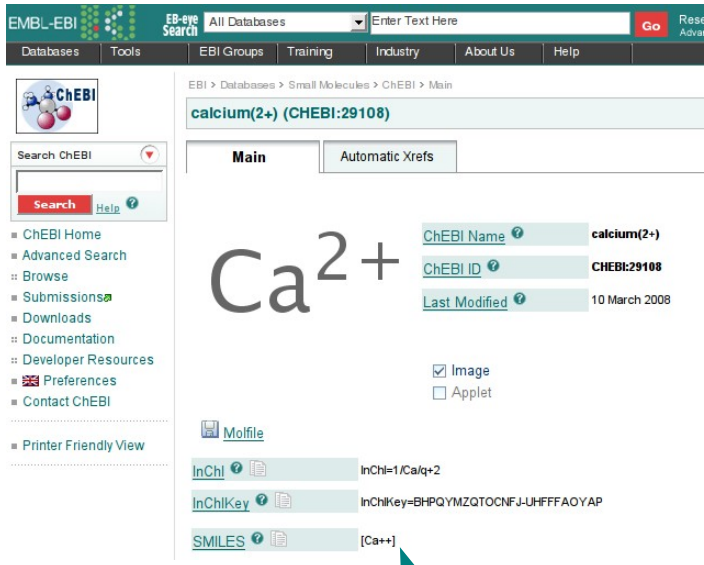
contractile proteins

Linking SBGN maps to external information

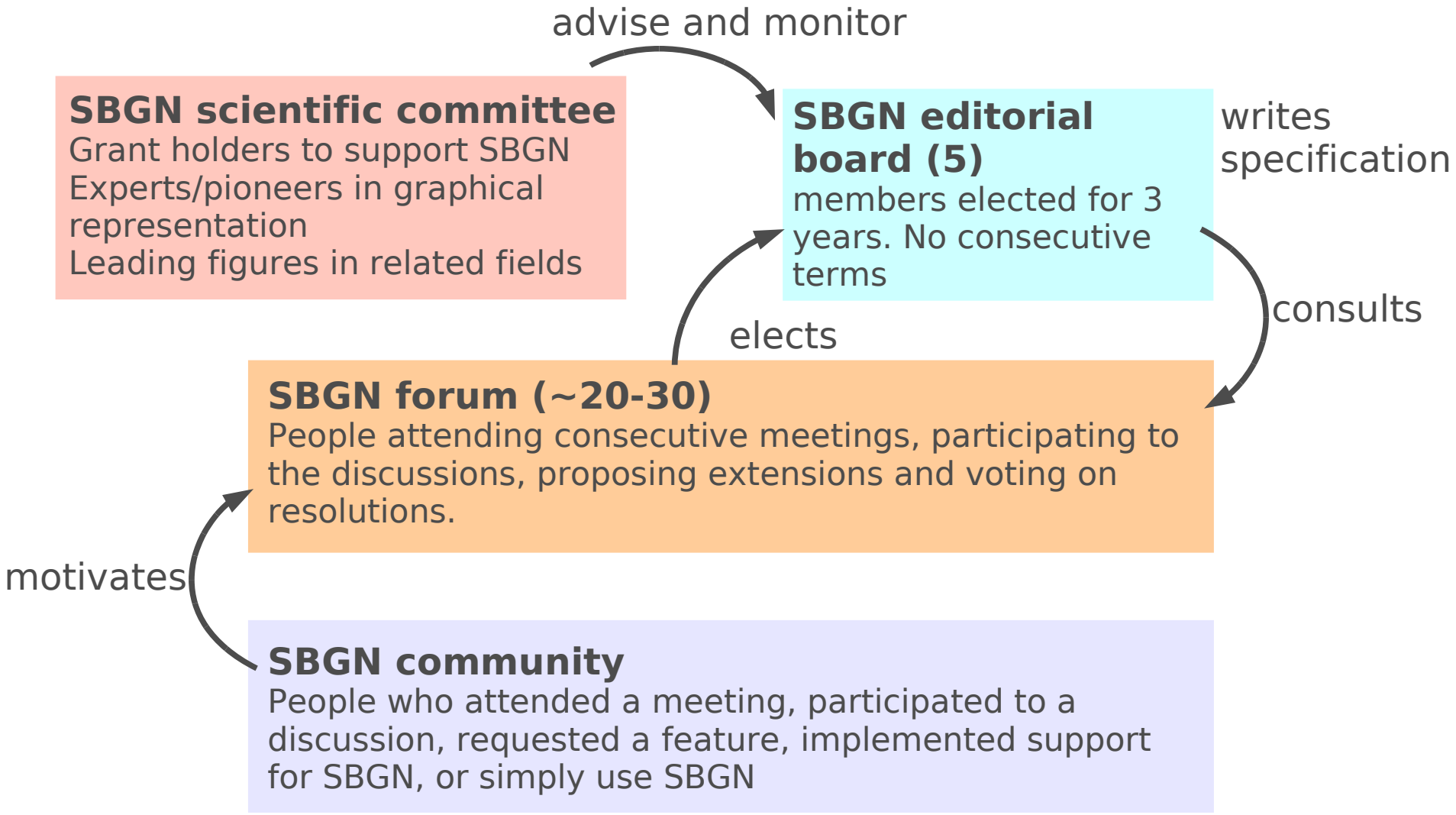


Linking SBGN maps to external information





Current structure of the SBGN community



Lack of biological semantics in SBML

```
<listOfCompartments>
  <compartment id="C">
</listOfCompartments>
<listOfSpecies>
  <species id="A"/>
  <species id="B"/>
  <species id="C"/>
</listOfSpecies>
<listOfReactions>
  <reaction>
    <listOfReactants>
      <speciesReference species="A"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C"/>
    </listOfModifiers>
    <kineticLaw>
      <math></math>
      <listOfParameters>
        <parameter id="U"/>
        <parameter id="V"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

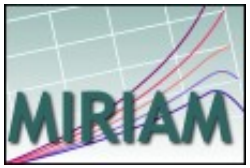
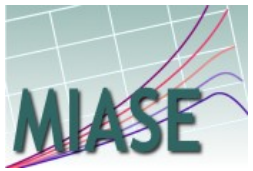
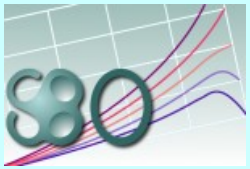
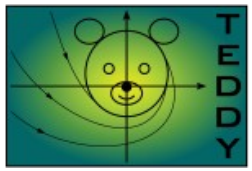
What are those?

("who" has been answered by MIRIAM annotations)

Do those affect/are affected by the reaction?

How should-I understand this?

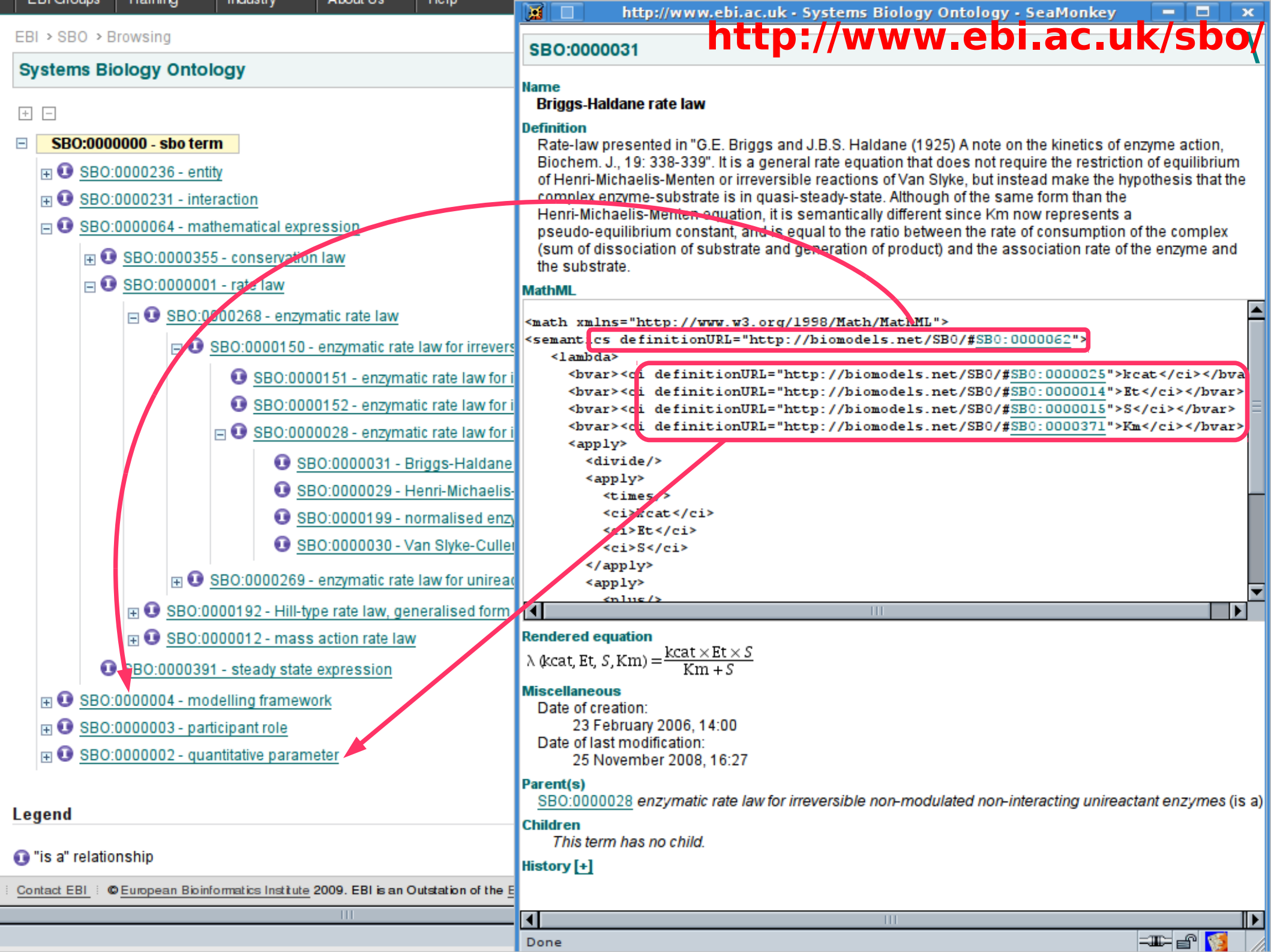
Model semantics

	Models	Simulation	Results
Minimal requirements			
Data-models		SED-ML	SBRML
Ontologies			

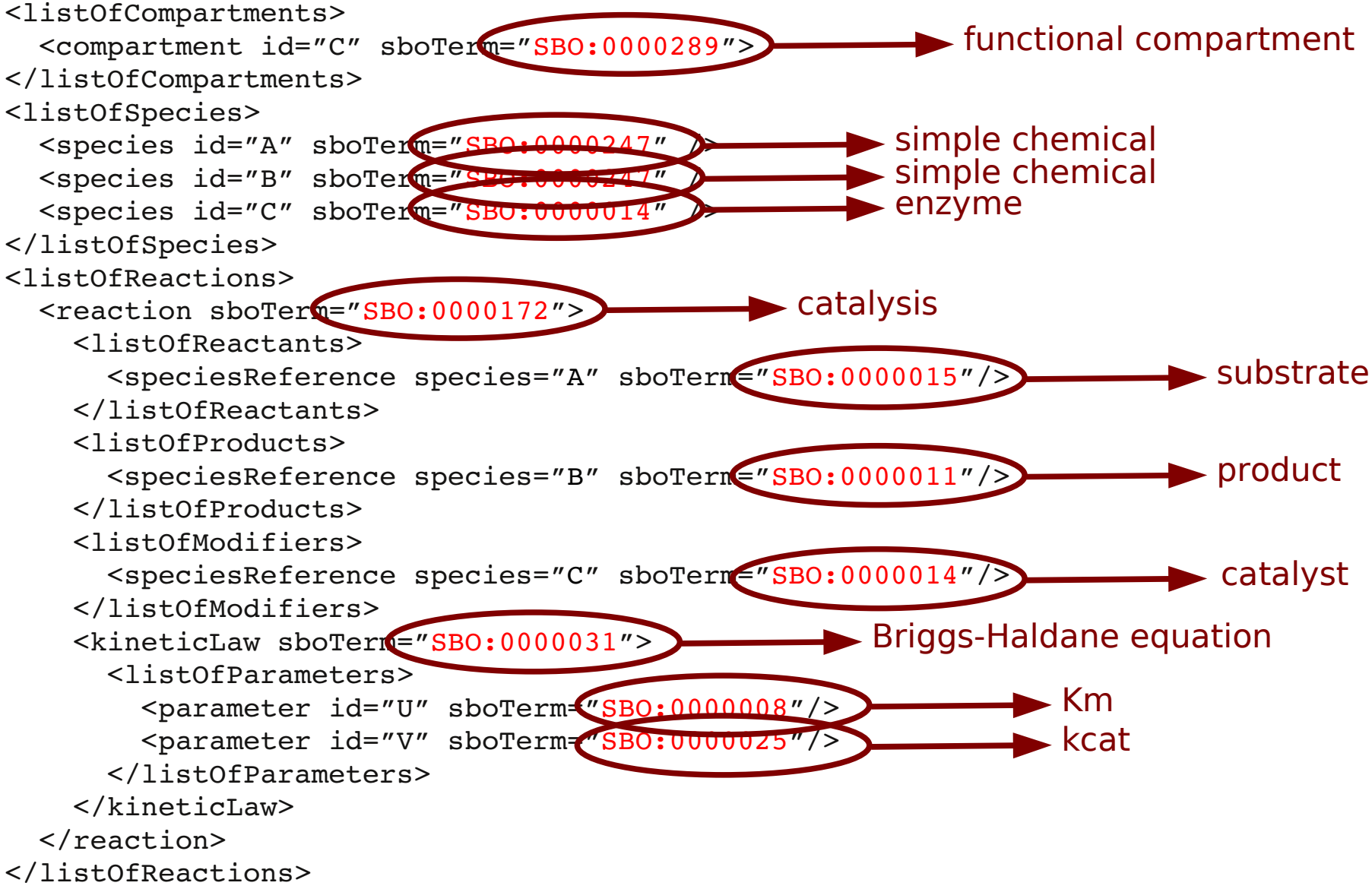
Systems Biology Ontology vocabularies

- Entities, that is existing objects, whether functional or material, such as "macromolecule", or "channel".
- Roles of reaction participants, including terms like "substrate", "catalyst" etc.
- Parameter used in quantitative models. This vocabulary includes terms like "Michaelis constant", "forward unimolecular rate constant"etc. A term may contain a precise mathematical expression stored as a MathML lambda function. The variables refer to other parameters.
- Mathematical expressions. Examples of terms are "mass action kinetics", "Henri-Michaelis-Menten equation" etc. A term may contain a precise mathematical expression stored as a MathML lambda function. The variables refer to the other vocabularies.
- Modelling framework to precise how to interpret the rate-law. E.g. "continuous modelling", "discrete modelling" etc.
- Event type, such as "catalysis" or "addition of a chemical group".

Le Novère N., Courtot M., Laibe C. Adding semantics in kinetics models of biochemical pathways. *Proc 2nd Intl Symp Experimental Standard Conditions of Enzyme Characterizations* (2007), 137-153. Available at http://www.beilstein-institut.de/index.php?id=196


EBI > SBO > Browsing
Systems Biology Ontology
SBO:0000000 - sbo term
SBO:0000236 - entity
SBO:0000231 - interaction
SBO:0000064 - mathematical expression
SBO:0000355 - conservation law
SBO:0000001 - rate law
SBO:0000268 - enzymatic rate law
SBO:0000150 - enzymatic rate law for irrevers
SBO:0000151 - enzymatic rate law for i
SBO:0000152 - enzymatic rate law for i
SBO:0000028 - enzymatic rate law for i
SBO:0000031 - Briggs-Haldane
SBO:0000029 - Henri-Michaelis-
SBO:0000199 - normalised enz
SBO:0000030 - Van Slyke-Culle
SBO:0000269 - enzymatic rate law for unireac
SBO:0000192 - Hill-type rate law, generalised form
SBO:0000012 - mass action rate law
SBO:0000391 - steady state expression
SBO:0000004 - modelling framework
SBO:0000003 - participant role
SBO:0000002 - quantitative parameter
Legend
"is a" relationship
Contact EBI
© European Bioinformatics Institute 2009. EBI is an Outstation of the E
http://www.ebi.ac.uk - Systems Biology Ontology - SeaMonkey
http://www.ebi.ac.uk/sbo/
SBO:0000031
Name
Briggs-Haldane rate law
Definition
Rate-law presented in "G.E. Briggs and J.B.S. Haldane (1925) A note on the kinetics of enzyme action, Biochem. J., 19: 338-339". It is a general rate equation that does not require the restriction of equilibrium of Henri-Michaelis-Menten or irreversible reactions of Van Slyke, but instead make the hypothesis that the complex enzyme-substrate is in quasi-steady-state. Although of the same form than the Henri-Michaelis-Menten equation, it is semantically different since Km now represents a pseudo-equilibrium constant, and is equal to the ratio between the rate of consumption of the complex (sum of dissociation of substrate and generation of product) and the association rate of the enzyme and the substrate.
MathML
<math xmlns="http://www.w3.org/1998/Math/MathML">
<semantics definitionURL="http://biomodels.net/SBO/#SBO:0000062">
<lambda>
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000025">kcat</ci></bva
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000014">Et</ci></bvar>
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000015">S</ci></bvar>
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000371">Km</ci></bvar>
<apply>
<divide/>
<apply>
<times/>
<ci>kcat</ci>
<ci>Et</ci>
<ci>S</ci>
</apply>
<apply>
Rendered equation
$\lambda(\mathrm{kcat}, \mathrm{Et}, S, \mathrm{Km}) = \frac{\mathrm{kcat} \times \mathrm{Et} \times S}{\mathrm{Km} + S}$
Miscellaneous
Date of creation:
23 February 2006, 14:00
Date of last modification:
25 November 2008, 16:27
Parent(s)
SBO:0000028 enzymatic rate law for irreversible non-modulated non-interacting unireactant enzymes (is a)
Children
This term has no child.
History [+]
Done

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" /
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```


<listOfCompartments>
<compartment id="C" sboTerm="SBO:0000289">
functional compartment
</listOfCompartments>
<listOfSpecies>
<species id="A" sboTerm="SBO:0000247" />
simple chemical
<species id="B" sboTerm="SBO:0000247" />
simple chemical
<species id="C" sboTerm="SBO:0000014" />
enzyme
</listOfSpecies>
<listOfReactions>
<reaction sboTerm="SBO:0000172">
catalysis
<listOfReactants>
<speciesReference species="A" sboTerm="SBO:0000015"/>
substrate
</listOfReactants>
<listOfProducts>
<speciesReference species="B" sboTerm="SBO:0000011"/>
product
</listOfProducts>
<listOfModifiers>
<speciesReference species="C" sboTerm="SBO:0000014"/>
catalyst
</listOfModifiers>
<kineticLaw sboTerm="SBO:0000031">
Briggs-Haldane equation
<listOfParameters>
<parameter id="U" sboTerm="SBO:0000008"/>
Km
<parameter id="V" sboTerm="SBO:0000025"/>
kcat
</listOfParameters>
</kineticLaw>
</reaction>
</listOfReactions>

Conversion between modeling frameworks

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" /
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

discrete simulator

$$v1 = \frac{(k_{-1} + V)}{U} \cdot [A] \cdot [C]$$

$$v2 = k_{-1} \cdot [D]$$

$$v3 = V \cdot [D]$$

continuous simulator

$$\nu = \frac{C \cdot V \cdot [A]}{(U + [A])}$$

SBML to BioPAX conversion using SBO

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" />
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

GO annotation

Small molecule

Small molecule

Protein

catalysis

physicalEntityParticipant

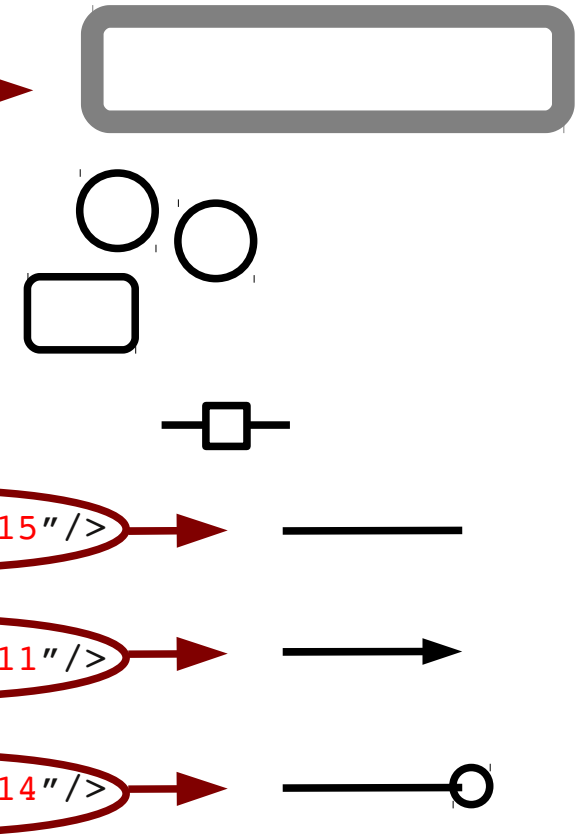
physicalEntityParticipant

physicalEntityParticipant

http://www.biopax.org/

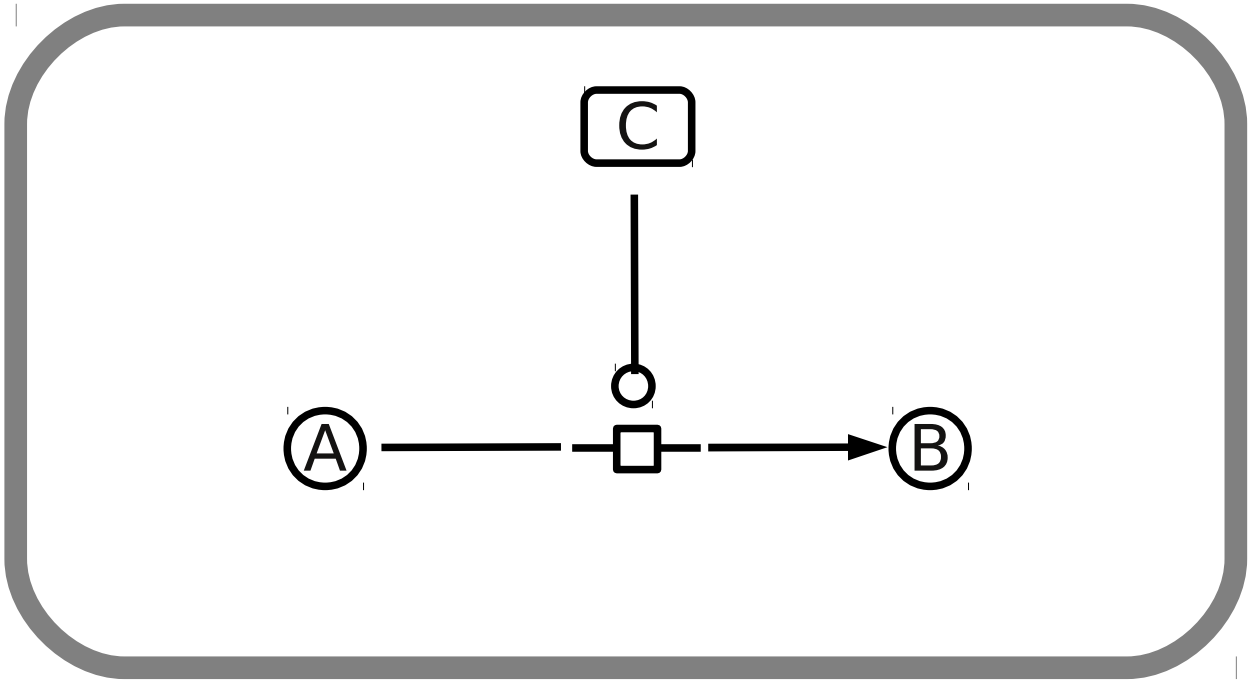
SBML to SBGN conversion using SBO

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" />
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

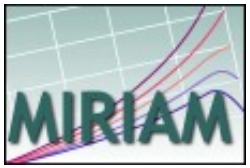
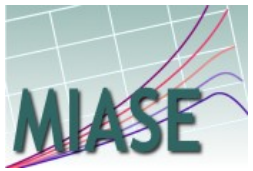
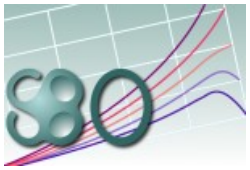
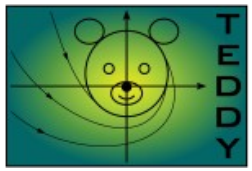


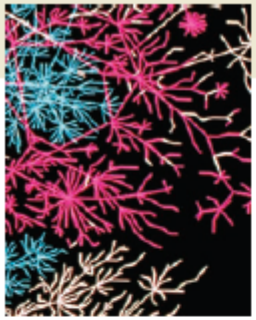
http://www.sbgn.org/

SBML to SBGN conversion using SBO



Simulation description

	Models	Simulation	Results
Minimal requirements			
Data-models		SED-ML	SBRML
Ontologies			



_computational BIOLOGY

PERSPECTIVE

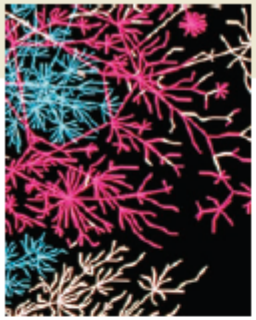
Publishing Group http://www.nature.com/naturebiotechnology

Minimum information requested in the annotation of biochemical models (MIRIAM)

Nicolas Le Novère[1,15], Andrew Finney[2,15], Michael Hucka[3], Upinder S Bhalla[4], Fabien Campagne[5], Julio Collado-Vides[6], Edmund J Crampin[7], Matt Halstead[7], Edda Klipp[8], Pedro Mendes[9], Poul Nielsen[7], Herbert Sauro[10], Bruce Shapiro[11], Jacky L Snoep[12], Hugh D Spence[13] & Barry L Wanner[14]

Most of the published quantitative models in biology are lost for the community because they are either not made available or they are insufficiently characterized to allow them to be reused. The lack of a standard description format, lack of stringent reviewing and authors' carelessness are the main causes for incomplete model descriptions. With today's increased interest in detailed biochemical models, it is necessary to define a minimum quality standard for the encoding of those models. We propose a set of rules for curating quantitative models of biological systems. These rules define procedures for encoding and annotating models represented in machine-readable form. We believe their

During the genomic era we have witnessed a vast increase in availability of large amounts of quantitative data. This is motivating a shift in the focus of molecular and cellular research from qualitative descriptions of biochemical interactions towards the quantification of such interactions and their dynamics. One of the tenets of systems biology is the use of quantitative models (see **Box 1** for definitions) as a mechanism for capturing precise hypotheses and making predictions[1,2]. Many specialized models exist that attempt to explain aspects of the cellular machinery. However, as has happened with other types of biological information, such as sequences, macromolecular structures or



_computational BIOLOGY

PERSPECTIVE

Publishing Group http://www.nature.com/naturebiotechnology

Minimur tion of biochem

Nicolas Le Novère
Julio Collado-Vid en[7],
Herbert Sauro[10],

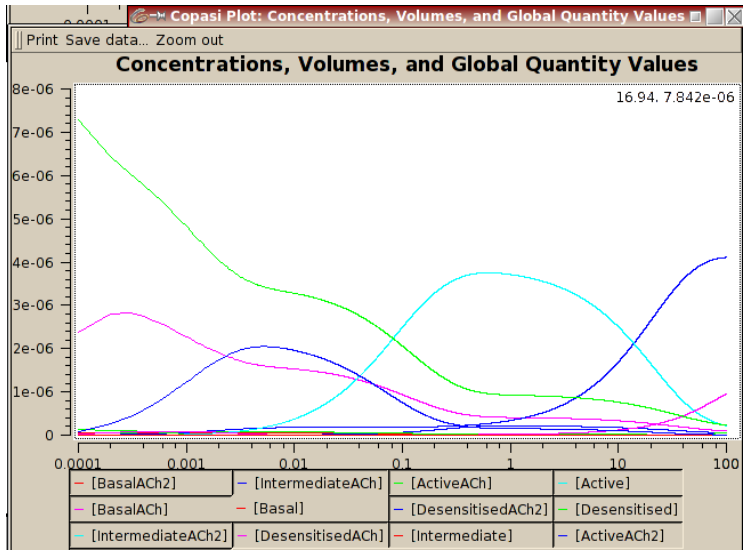
6. The model, when instantiated within a suitable simulation environment, must be able to reproduce all relevant results given in the reference description that can readily be simulated. Not only does the simulation have to provide results qualitatively similar to the reference description, such as oscillation, bistability, chaos, but the quantitative values of variables, and their relationships (e.g., the shape of the phase portrait) must be reproduced within some epsilon, the difference being attributable to the algorithms used to run the simulation, and the

Most of the published lost for the community because they are either not made available or they are insufficiently characterized to allow them to be reused. The lack of a standard description format, lack of stringent reviewing and authors' carelessness are the main causes for incomplete model descriptions. With today's increased interest in detailed biochemical models, it is necessary to define a minimum quality standard for the encoding of those models. We propose a set of rules for curating quantitative models of biological systems. These rules define procedures for encoding and annotating models represented in machine-readable form. We believe their

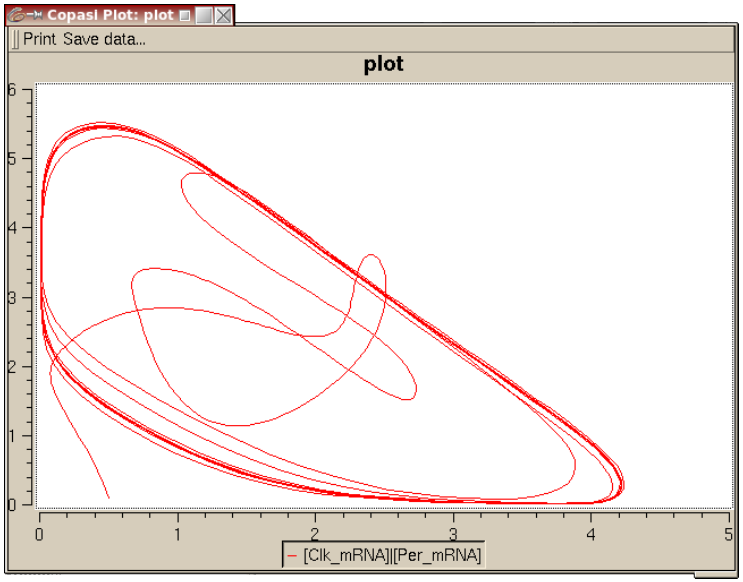
increase in availability of large amounts of quantitative data. This is motivating a shift in the focus of molecular and cellular research from qualitative descriptions of biochemical interactions towards the quantification of such interactions and their dynamics. One of the tenets of systems biology is the use of quantitative models (see **Box 1** for definitions) as a mechanism for capturing precise hypotheses and making predictions[1,2]. Many specialized models exist that attempt to explain aspects of the cellular machinery. However, as has happened with other types of biological information, such as sequences, macromolecular structures or

Reproduction of published simulation results

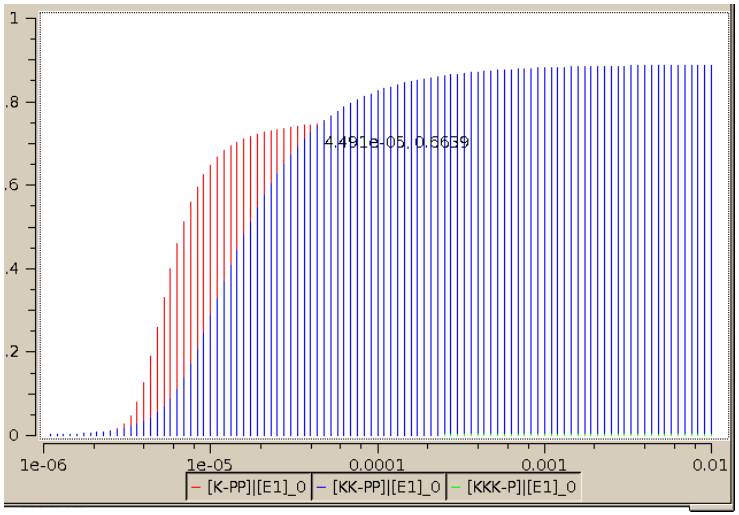
Edelstein et al 1996 (BIOMD0000000002)



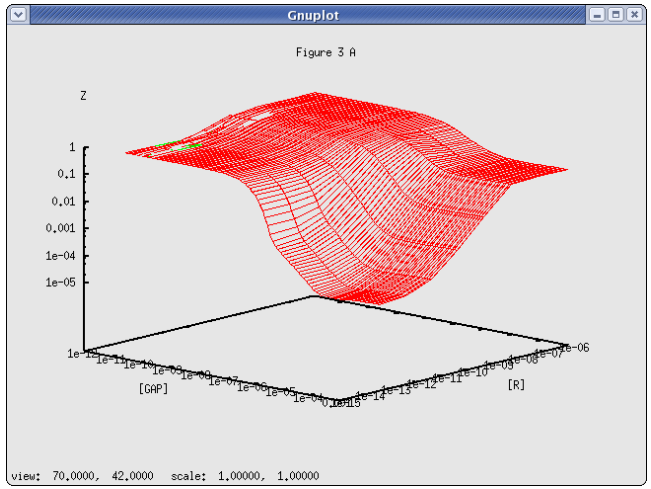
Ueda, Hagiwara, Kitanol 2001 (BIOMD0000000022)



Huang & Ferrell (BIOMD0000000009)



Bornheimer et al 2004 (BIOMD0000000086)



Waltemath D., Adams R., Beard D.A., Bergmann F.T., Bhalla U.S, Britten R., Chelliah V., Cooling M.T., Cooper J., Crampin E., Garny A., Hoops S., Hucka M., Hunter P., Klipp E., Laibe C., Miller A., Moraru i., Nickerson D., Nielsen P., Nikolski M., Sahle S., Sauro H., Schmidt H., Snoep J.L., Tolle D., Wolkenhauer O., Le Novère N.

Minimum Information About a Simulation Experiment (MIASE) *In Revision*

Köhn D.,Le Novère N. SED-ML

An XML Format for the Implementation of the MIASE Guidelines.

Proc 6th Conf Comput Meth Syst Biol (2008), Heiner M and Uhrmacher AM eds, *Lecture Notes in Bioinformatics*, 5307: 176-190.

Description of models

```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```

Description of models

```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```

Any model description in XML such as SBML, CellML VCML etc.

Description of models

```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```

Description of models

Data type: ***BioModels Database***

General | Tags | Annotation

General information about the data type

Name		
Identifier		MIR:00000007
Name		BioModels Database
Synonyms		BioModels
URIs		
Official URN		urn:miriam:biomodels.db
Deprecated		http://www.ebi.ac.uk/biomodel
Definition		BioModels Database is a data models of biological interests.
Identifier Pattern		^(BIOMD|MODEL)\d{10}$
Resource #1	**Data Entry**	http://www.ebi.ac.uk/biomodel
	Data Resource	http://www.ebi.ac.uk/biomodel
	Information	BioModels, a Database of Ann
	Institution	European Bioinformatics Insti
Resource #2	**Data Entry**	http://biomodels.caltech.edu/$
	Data Resource	http://biomodels.caltech.edu/
	Information	Mirror of BioModels Database
	Institution	California Institute of Technol
URL(s)		http://srs.ebi.ac.uk/srsbin/c
Date of creation		2006-08-14 19:38:06 GMT
Date of last modification		2009-06-05 17:43:08 GMT

Go back to the list of data types

BioModels Home | Browse Models | Submit | Sign In | Support | About | Search

BIOMD0000000021 - Leloup1999_CircClock

SBML L2 V1 | Other formats | Actions | Submit Model Comment/Bug

Overview | Model | Math | Physical entities | Parameters

Reference Publication

Publication ID: 10366496

J Theor Biol 1999 Jun;198(3):445-59.
Chaos and birhythmicity in a model for circadian oscillations of the PER and TIM proteins in drosophila
Leloup JC, Goldbeter A.
Unite de Chronobiologie Theorique, Faculte des Sciences, Universite Libre de Bruxelles, Campus Plaine, C.P. 231, B-1050 Brussels, Belgium. [more]

Model

Original Model: *Unspecified*

Submitter: Nicolas Le Novère

Submission Date: 2005-09-13T13:24:15+00:00

Last Modification Date: 2007-09-25T09:32:00+00:00

Creation Date: 2005-06-29T10:27:52+00:00

Creators: Nicolas Le Novère
Bruce Shapiro

Result:

set #1	bqbiol:is	**Taxonomy** Drosophila melanogaster **KEGG Pathway** dme04710
set #2	bqbiol:isVersionOf	**Gene Ontology** circadian rhythm

Notes

Terms of Use | EBI Funding | Contact EBI | © European Bioinformatics Institute 2006-2008. EBI is an Outstation of the European Molecular Biology Laboratory.

Description of models

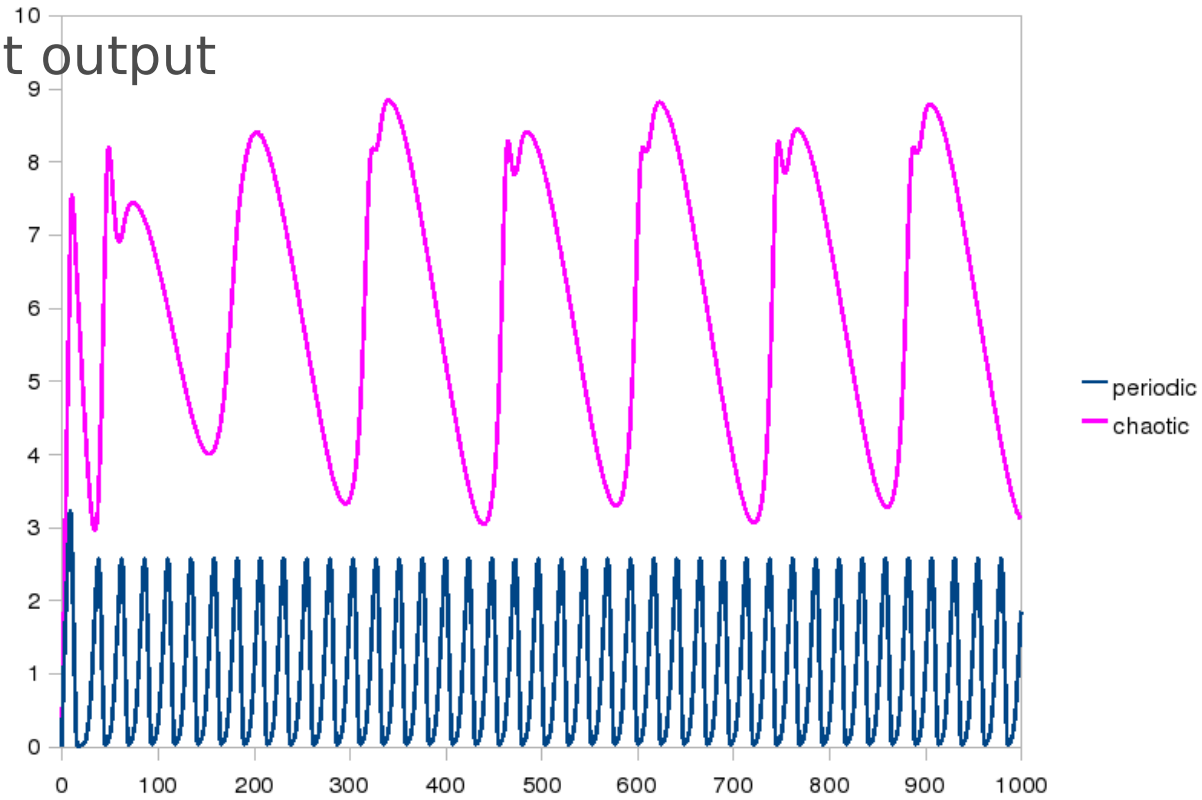
```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```

Description of models

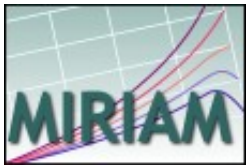
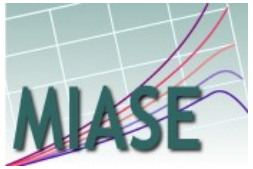
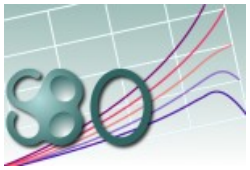
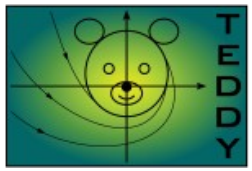
```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```

Production of results

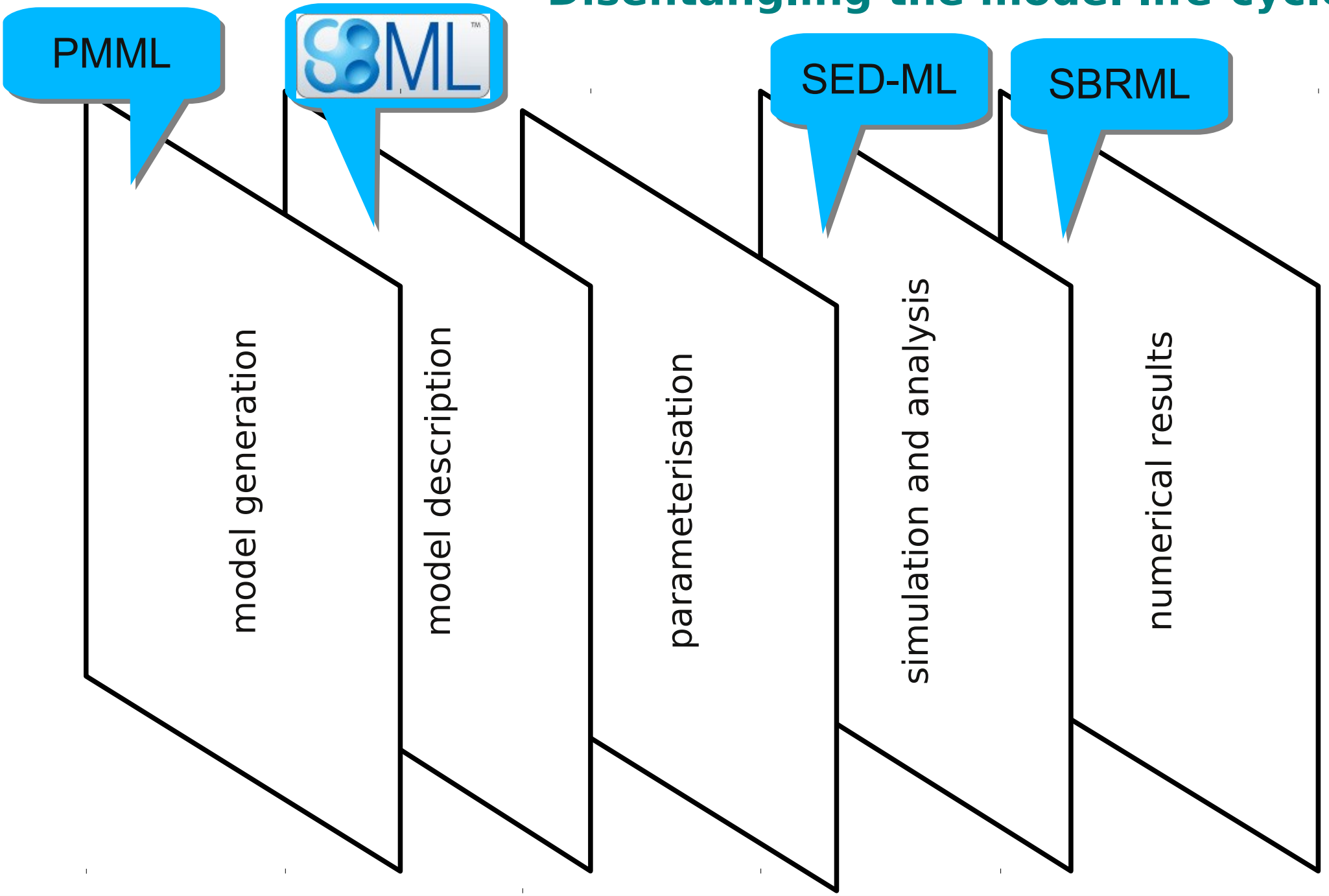
- Description of models
- Description of simulations
- Description of tasks
- Description of result post-processing
- Description of experiment output



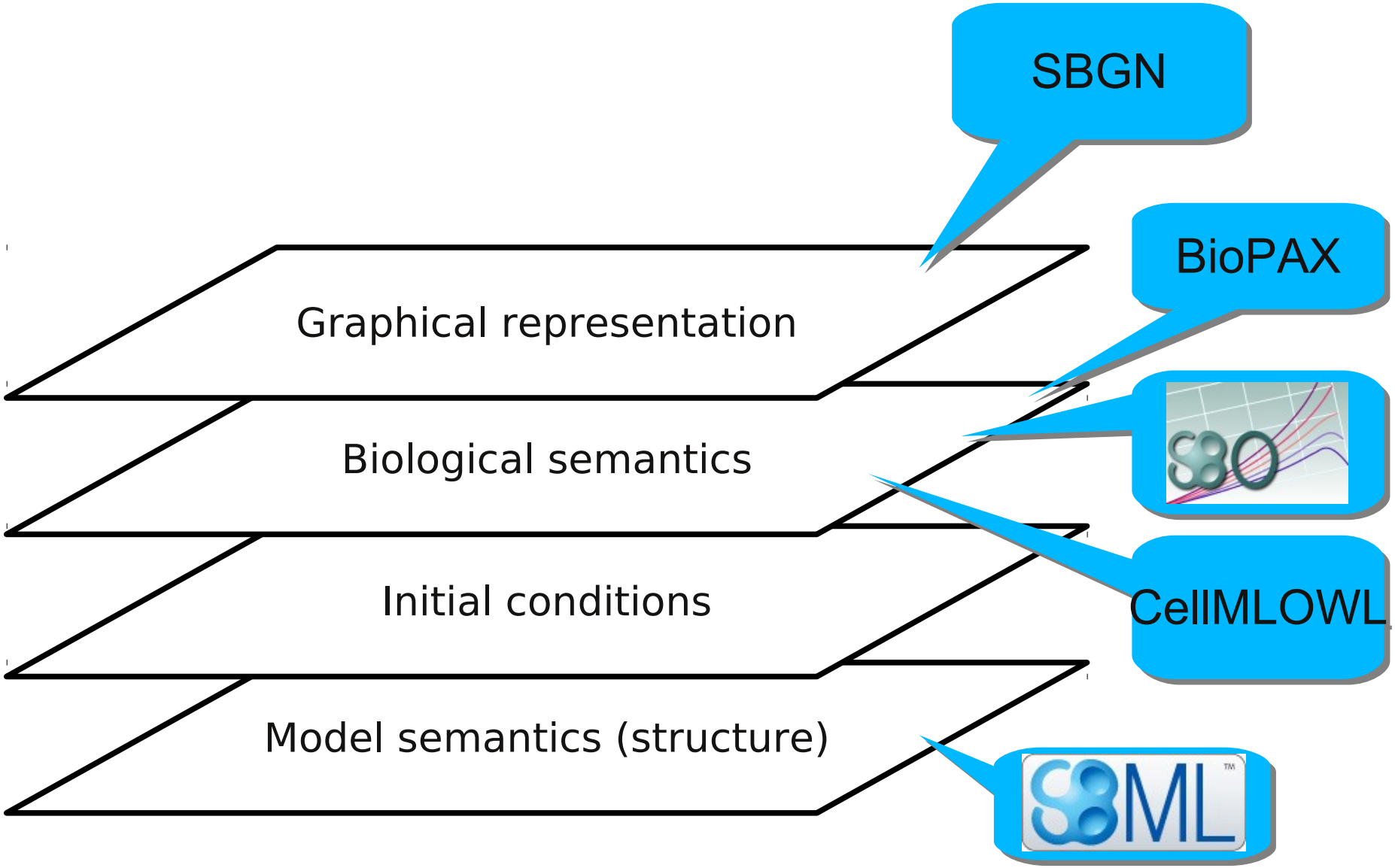
Is the mosaic of standards really complete?

	Models	Simulation	Results
Minimal requirements			
Data-models		SED-ML	SBRML
Ontologies			

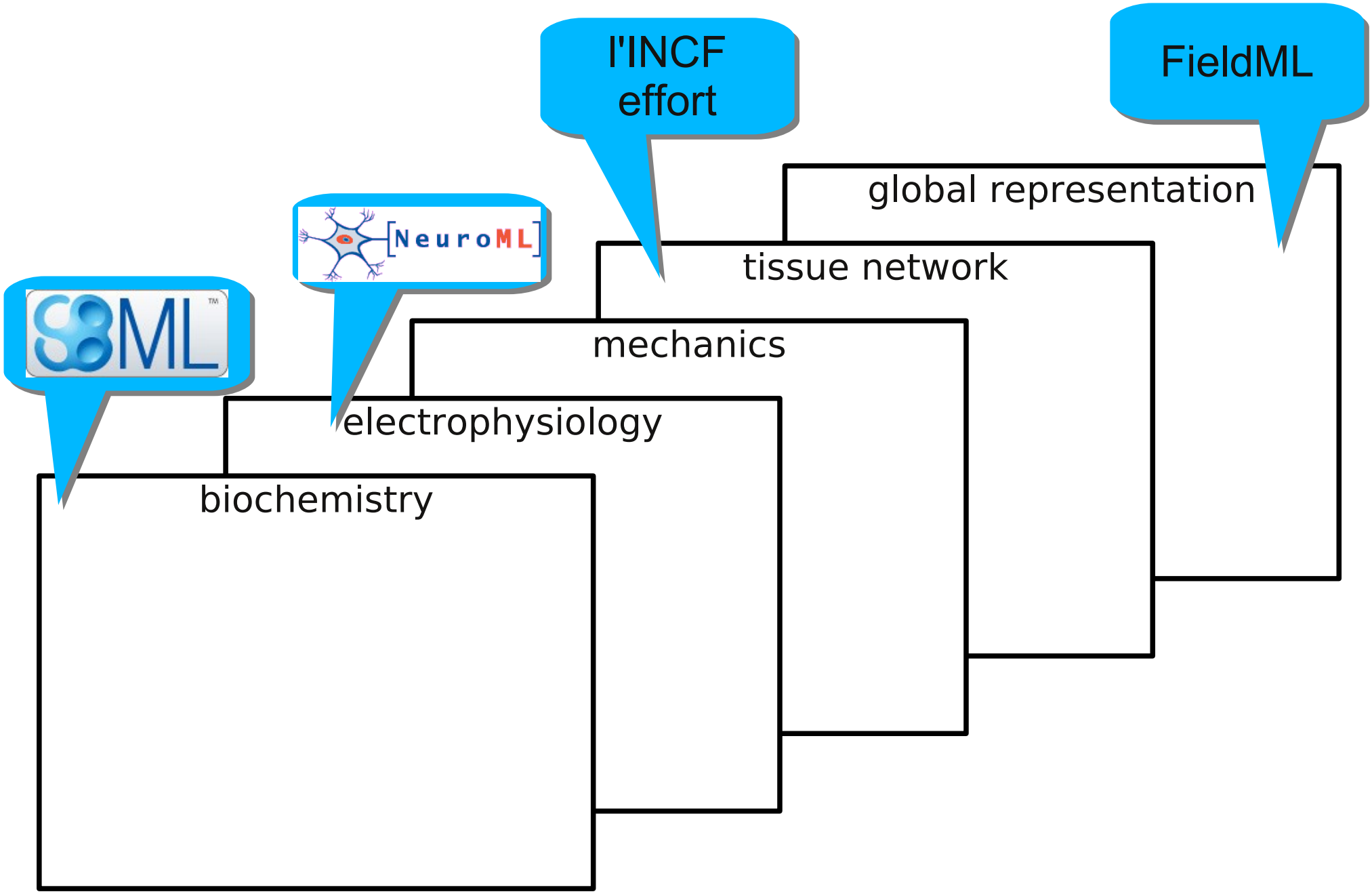
Disentangling the model life-cycle



Disentangling the level of discourse



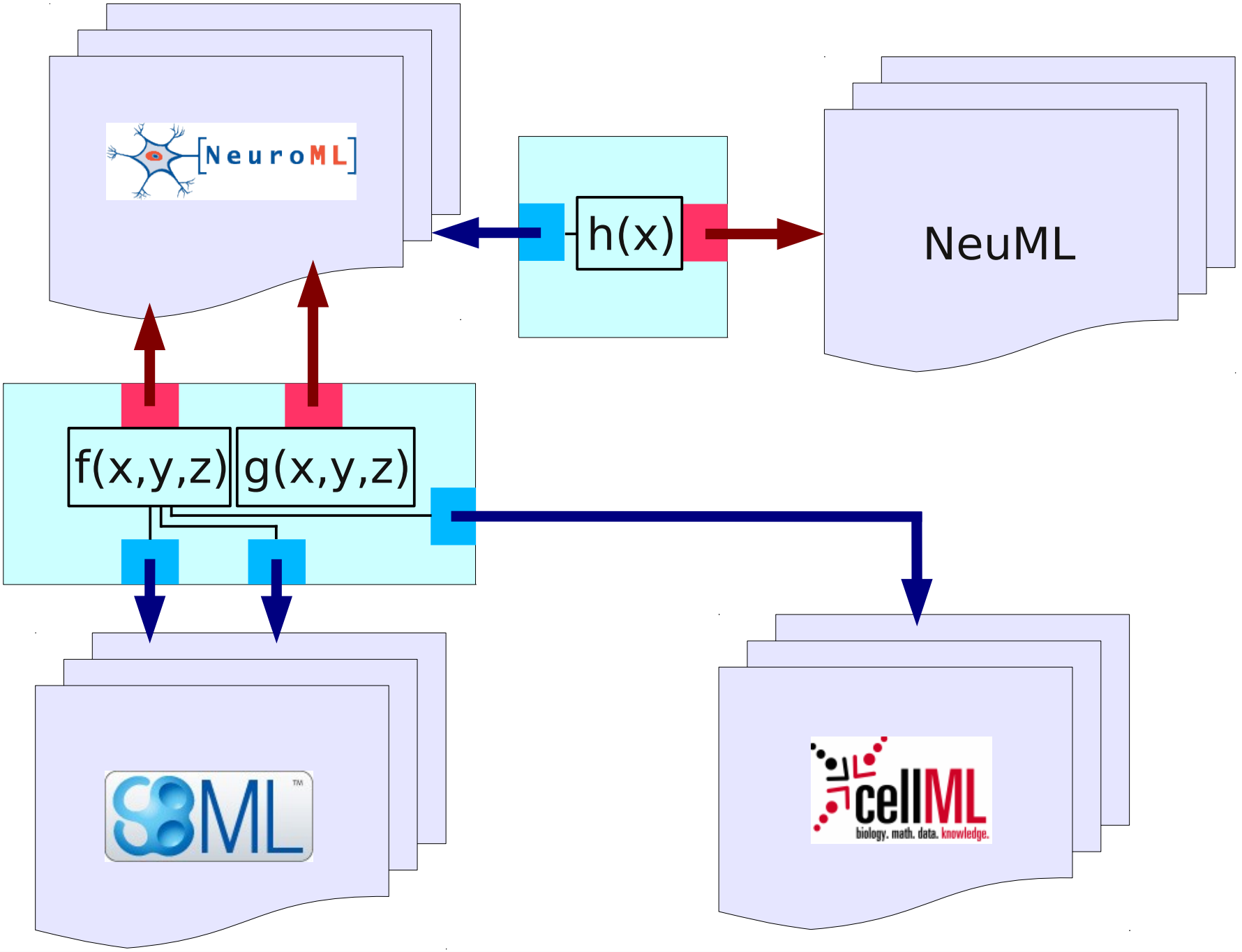
Non-overlapping languages



Interfacing standards in the three dimensions

- SBML to BioPAX: interface using annotations (MIRIAM annotations and SBO terms), e.g.
  - mapping between Species and PhysicalEntity
  - mapping between Reactions and PhysicalInteraction
- Usage of SBML descriptions (or CellML or VCML) in SED-ML: Identification of variables using XPath
- Descriptions using SBML and NeuroML: Interface based on shared namespaces

Multi-scale representation using adapters



Threats to the whole enterprise

- Current efforts are entirely dependent on key people (SBML: Mike Hucka, CellML: Peter Hunter/Poul Nielsen, NeuroML: Padraig Gleeson, SBGN: NLN). Their disengagement means disaggregation.
- Current funding structure is fragile. Many different grants, sometimes only supporting meetings (SBGN), none of them infrastructure rolling funding, often tied to individuals.
- Current efforts are not immune against intellectual property claims that would destroy the community (e.g. Caltech and SBML)
- Existing standards are developed with very different approaches, quality checks, and are based on completely different assumptions (e.g. NeuroML assumes implicit knowledge)
- APIs needs industry-grade support, incompatible with standard academic usages and possibilities

STANDARDS

Web Design and Applications
Web Architecture
Semantic Web
XML Technology
Web of Services
Web of Devices
Browsers and Authoring Tools
All Standards and Drafts
About W3C Standards

WEB DESIGN AND APPLICATIONS

On this page → technology topics • news • upcoming events and talks

Web Design and Applications involve the standards for building and Rendering Web pages, including HTML, CSS, SVG, Ajax, and other technologies for Web Applications ("WebApps"). This section also includes information on how make pages accessible to people with disabilities (WCAG), internationalized, and work on mobile devices.

HTML & CSS

HTML and CSS are the fundamental technologies for building Web pages: HTML (html and xhtml) for structure, CSS for style and layout. Find resources for good Web page design as well as helpful tools.

Scripting and Ajax

Standard APIs for client-side Web Application development include those for Geolocation, XMLHttpRequest (Ajax), and mobile widgets. W3C standards for document models (the "DOM") and technologies such as XBL allow content providers to create interactive documents through scripting.

Graphics

W3C is the home of the widely deployed PNG raster format, SVG vector format, and the Canvas API. WebCGM is a more specialized format used, for example, in the fields of automotive engineering, aeronatics.

Audio and Video

Some of the W3C formats that enable authoring audio and video presentations include HTML, SVG, and SMIL (for synchronization). W3C is also working on a timed text format for captioning and other applications.

Accessibility

W3C's Web Accessibility Initiative (WAI) has published Web Content Accessibility Guidelines (WCAG) to help authors create content that is accessible to people with disabilities. WAI-ARIA gives authors more tools to create accessible Web Applications by providing additional semantics about widgets and behaviors.

Internationalization

W3C has a mission to design technology that works across cultures and languages. W3C standards such as HTML and XML are built on Unicode, for instance. In addition, W3C has published guidance for authors related to language tags bi-directional (bidi) text, and more.

Mobile Web

W3C promotes "One Web" that is available on any device. W3C's Mobile Web Best Practices help authors understand how to

Privacy

The Web is a powerful tool for communications and transactions of all sorts. It is important to consider privacy and security

Math on the Web

Mathematics and formula are used on the Web for business reports, education materials and scientific research. W3C's

Poster presented by Mike Hucka at the 6th ICSB Boston 2005

BioModels.net

Michael Hucka
Biological Network Modeling Center
California Institute of Technology
Pasadena, California 91125, USA

Andrew Finney
Physiomics PLC
Oxford Science Park
Oxford, OX4 4GA, UK

Nicolas Le Novère
European Bioinformatics Institute
Wellcome Trust Genome Campus, Hinxton
Cambridge, CB10 1SD, UK

For computational modeling to become more widely used in biology, researchers must be able to exchange and share their results. The development and broad acceptance of common representation formats such as SBML is a crucial step in that direction, allowing researchers to exchange and build upon each other's work with greater ease and accuracy.

BioModels.net is another step: an international effort to
1) define agreed-upon standards for model curation
2) define agreed-upon vocabularies for annotating models, and
3) provide a free, centralized, publicly-accessible database of annotated models in SBML and other structured formats.

SBO

The **Systems Biology Ontologies** (SBO) are a set of interlinked controlled vocabularies tailored to the needs of the systems biology modeling community. The aim is to capture consensus definitions of commonly-used concepts and their relationships. Using SBO terms as a basis for annotating models and performing database searches (e.g., in BioModels Database) can lead to a greater degree of consistency, quality and interoperability.

BioModels Database

BioModels Database is a free, public repository of published quantitative models of biochemical and cellular systems. The models are curated to verify that they correspond to the reference publication and give the proper numerical results. Curators annotate the components of the models with terms from controlled vocabularies and links to external relevant data resources, allowing users to search accurately for the models they need and precisely identify their components.

References

Le Novère N, Bornstein B, Broicher A, Courtot M, Donizelli M, Dharuri H, Li L, Sauro H, Schilstra M, Shapiro B, Snoep JL, Hucka M. BioModels Database: A Free, Centralized Database of Curated, Published, Quantitative Kinetic Models of Biochemical and Cellular Systems. (2006) *Nucleic Acids Res.*, in press.

Le Novère N, Finney A, Hucka M, Bhalla U, Campagne F, Collado-Vides J, Crampin EJ, Halstead M, Klipp E, Mendes P, Nielsen P, Sauro H, Shapiro B, Snoep JL, Spence HD, Wanner BL. Minimal Information Requested In the Annotation of biochemical Models (MIRIAM). (2005) *Nature Biotechnology*, in press.

MIRIAM

The **Minimum Information Requested in the Annotation of Models** (MIRIAM) is a proposed set of rules for curating quantitative models of biological systems. The rules define procedures for encoding and annotating models represented in machine-readable form. Their application will enable users to (a) have confidence that curated models are an accurate reflection of their associated reference description; (b) search collections of curated models with precision; (c) quickly identify the biological phenomena that a given curated model or model constituent represents; and (d) facilitate model reuse and composition into large subcellular models.

Acknowledgments

Funding for different aspects of the BioModels.net effort currently comes from the following organizations: the National Institute of General Medical Sciences (USA, grant GM070923-02S1), the International Joint Research Program of NEDO (Japan), DARPA BioComp Bio-SPICE (USA), the European Molecular Biology Laboratory (EU), the California Institute of Technology (USA), the Systems Biology Institute (Japan), and the University of Hertfordshire (UK).

BioModels.Net

BioModels.net

The Next Step After Standard Formats

For computational modeling to become more widely used in biological research, researchers must be able to exchange and share their results. The development and broad acceptance of common model representation formats such as **SBML** is a crucial step in that direction, allowing researchers to exchange and build upon each other's work with greater ease and accuracy.

The BioModels.net project is another step: an international effort to:

1. define agreed-upon standards for model curation
2. define agreed-upon vocabularies for annotating models with connections to biological data resources
3. provide a free, centralized, publicly-accessible database of annotated, computational models in SBML and other structured formats

Helping to Define Community Standards

To facilitate assembling useful collections of quantitative models of biological phenomena, it is crucial to establish standards for the vocabularies used in model annotations as well as criteria for minimum quality levels of those models. The BioModels.net project aims to bring together a community of interested researchers to address these issues. We are working towards defining these standards through white papers and process definitions. All of the products of our efforts are open and freely available through this site.

Standards and Processes Developed Hand-in-Hand with a New Database

The database component of BioModels.net is especially designed for working with *annotated* computational models: each model is carefully reviewed and augmented by human annotators on the BioModels.net team to add metadata linking the model elements to other biological databases and resources. The **BioModels Database** at the **EBI** system goes far beyond other collections of models by being a *true* database, featuring browsing, cross-referencing, searching, and facilities for visualization, exporting models in different formats, and remote API access.

Projects

The projects we are currently coordinating are:

- **BioModels Database**
- **MIRIAM** and the associated **set of qualifiers** and **MIRIAM Resources**
- **SBO**

Requirements for a global coordination structure

- What?
  - Set of interoperable description languages
  - Cover all aspects of modelling and simulation
  - Cover all type of descriptions / views of the real
  - Role of community-maintained ontologies.
- How?
  - Independence towards Institutions, funders and individuals
  - Able to receive funds, to employ staff
  - Role of European Research Infrastructures? (ELIXIR, ISBE)
- Who?
  - Community developing their standards: Systems Biology, Physiology, Neuroscience (INCF), Drug discovery
  - Other players in knowledge-representation (W3C, ...)
  - Academic and corporate users: Modeling platforms (Matworks ...), Pharma (Pistoia alliance) ...

A first step: common meetings

- January 2008: SBGN hackathon; BioModels DB; MIRIAM; SBO
- March 2009: SBML hackathon; BioModels DB; MIRIAM; SBO
- April 2009: CellML; SED-ML; SBGN hackathon
- May 2010: SBML hackathon; BioModels DB; MIRIAM; SBO; SED-ML
- October 2010: 1st ***COMBINE*** MEETING with SBML; SBGN; ??
- From now on, two grouped annual meetings
  - ***COMBINE*** forum: presentation of support, discussion about future developments and collaboration etc.
  - ***HARMONY*** hackathon: developing support, writing specifications, tinkering with interoperability etc.

Acknowledgements

Visionnary: **Hiroaki Kitano**

SBML editors: Frank Bergmann, *Andrew Finney, Stefan Hoops*, ***Michael Hucka***, *Nicolas Le Novère, Sarah Keating*, Sven Sahle, *Herbert Sauro*, Jim Schaff, Lucian Smith, Darren Wilkinson

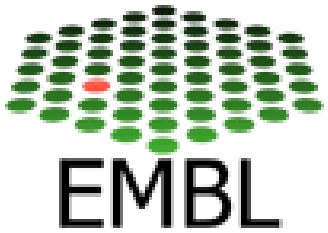
SBGN editors: Emek Demir, *Michael Hucka*, Nicolas Le Novère, Huaiyu Mi, Stuart Moodie, Falk Schreiber, *Anatoly Sorokin*

SBO/MIRIAM: Mélanie Courtot, Nick Juty, Camille Laibe

SED-ML/MIASE: Dagmar Köhn/Waltemath

The EBI group Computational Systems Neurobiology

The whole community of Computational Systems Biology



THE END

THE BEGINNING